

Oxford
LINGUISTICS

parameters
diffusion
transparency
principle

Diachronic Syntax

Ian Roberts

variation
DIRECTION
reanalysis

OXFORD TEXTBOOKS IN LINGUISTICS

OXFORD TEXTBOOKS IN LINGUISTICS

Diachronic Syntax

OXFORD TEXTBOOKS IN LINGUISTICS

General editors: **Keith Brown**, University of Cambridge; **Eve V. Clark**, Stanford University;
April McMahon, University of Edinburgh; **Jim Miller**, University of Auckland;
Lesley Milroy, University of Michigan

This series provides lively and authoritative introductions to the approaches, methods, and theories associated with the main subfields of linguistics.

The Grammar of Words
An Introduction to Linguistic Morphology
by Geert Booij

A Practical Introduction to Phonetics
Second edition
by J. C. Catford

Meaning in Language
An Introduction to Semantics and Pragmatics
Second edition
by Alan Cruse

Principles and Parameters
An Introduction to Syntactic Theory
by Peter W. Culicover

A Semantic Approach to English Grammar
by R. M. W. Dixon

Semantic Analysis
A Practical Introduction
by Cliff Goddard

Pragmatics
by Yan Huang

Diachronic Syntax
by Ian Roberts

Cognitive Grammar
An Introduction
by John R. Taylor

Linguistic Categorization
Third edition
by John R. Taylor

Diachronic Syntax

Ian Roberts

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Ian Roberts 2007

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2007

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed by Biddles Ltd., www.biddles.co.uk.

ISBN: 019-928366-4 978-019-928366-8 hbk

ISBN: 019-925398-6 978-019-925398-2 pbk

Contents

Preface xi

List of Abbreviations and Acronyms xii

Introduction 1

Further reading 8

1. Comparative and historical syntax in the principles-and parameters-approach 11

Introduction 11

1.1. UG and variation in grammatical systems 19

1.2. The null-subject parameter 24

1.2.1. The null-subject parameter in the synchronic dimension 24

1.2.2. The null-subject parameter in the diachronic dimension: changes in the history of French 33

1.3. Verb-movement parameters 40

1.3.1. Verb-movement in the synchronic dimension 41

1.3.1.1. Verb-movement to T 41

1.3.1.2. V-movement to C:
full and residual V2 48

1.3.1.3. Further properties related to
verb-movement 55

1.3.2. Verb-movement in the diachronic dimension 56

1.3.2.1. V-to-T in earlier English 56

1.3.2.2. V2 in diachrony 58

1.4. Negative concord 64

1.4.1. Negative concord synchronically 64

1.4.2. Negative concord in the diachronic dimension:
the development of French n-words 77

1.5. Wh-movement 81

1.5.1. The wh-movement parameter 81

1.5.2. Wh-movement in the diachronic domain:
Old Japanese 90

- 1.6. Head-complement order 92
 - 1.6.1. Head-complement order synchronically 92
 - 1.6.2. Head-complement order diachronically 102
- 1.7. Summary 108
 - Further reading 110
- 2. **Types of syntactic change** 121
 - Introduction 121
 - 2.1. Reanalysis 122
 - 2.1.1. The nature of reanalysis 122
 - 2.1.2. The Transparency Principle 127
 - 2.1.3. Phonology and reanalysis 129
 - 2.1.4. Expressing parameters 132
 - 2.1.5. Reanalysis and the poverty of the stimulus 140
 - 2.1.6. Conclusion 141
 - 2.2. Grammaticalization 141
 - 2.3. Argument structure 149
 - 2.3.1. Thematic roles and grammatical functions 149
 - 2.3.2. Changes in English psych verbs and recipient passives 152
 - 2.4. Changes in complementation 161
 - 2.5. Word-order change: OV > VO in English 175
 - 2.5.1. Introduction 175
 - 2.5.2. Early typological approaches to word-order change 176
 - 2.5.3. Generative accounts and directionality parameters 180
 - 2.5.4. 'Antisymmetric' approaches to word-order change 189
 - 2.5.5. Conclusion 197
 - 2.6. Conclusion to Chapter 2 198
 - Further reading 198
- 3. **Acquisition, learnability, and syntactic change** 207
 - Introduction 207
 - 3.1. First-language acquisition from a principles-and-parameters perspective 209
 - 3.2. The logical problem of language change 226
 - 3.3. The changing trigger 236

-
- 3.3.1. Contact-driven parameter-resetting 236
 - 3.3.2. Cue-driven parameter-resetting 242
 - 3.3.3. Morphologically-driven parameter-resetting 245
 - 3.3.4. Conclusion 251
 - 3.4. Markedness and complexity 251
 - 3.4.1. The concept of markedness 251
 - 3.4.2. Markedness and parameters 253
 - 3.4.3. The Subset Principle 256
 - 3.4.4. Markedness and core grammar 261
 - 3.4.5. Markedness and inflectional morphology 261
 - 3.4.6. Markedness, directionality,
and uniformitarianism 264
 - 3.4.7. Conclusion 266
 - 3.5. Parameter setting and change 266
 - 3.5.1. A format for parameters 267
 - 3.5.2. A markedness convention for syntax 272
 - 3.5.3. From unmarked to marked 275
 - 3.5.4. Networks of parameters 277
 - 3.5.5. Conclusion 282
 - 3.6. Conclusion to Chapter 3 282
 - Further reading 284
 - 4. The dynamics of syntactic change 291**
 - Introduction 291
 - 4.1. Gradualness 293
 - 4.1.1. Introduction 293
 - 4.1.2. Lexical diffusion 297
 - 4.1.3. Microparametric change 300
 - 4.1.4. Formal optionality 305
 - 4.1.5. The Constant Rate Effect 309
 - 4.1.6. Conclusion 314
 - 4.2. The spread of syntactic change 315
 - 4.2.1. Introduction 315
 - 4.2.2. Orderly differentiation
and social stratification 316
 - 4.2.3. Grammars in competition 319
 - 4.2.4. Formal optionality again 331
 - 4.2.5. Abduction and actuation 333

- 4.2.6. Change in progress? Null subjects in Brazilian Portuguese 335
 - 4.2.7. Conclusion 339
 - 4.3. Drift: the question of the direction of change 340
 - 4.3.1. Introduction 340
 - 4.3.2. Typological approaches to drift 342
 - 4.3.3. Drift and parametric change 345
 - 4.3.4. Cascading parameter changes in the history of English 351
 - 4.3.5. Conclusion 357
 - 4.4. Reconstruction 357
 - 4.4.1. Introduction 357
 - 4.4.2. Traditional comparative reconstruction 358
 - 4.4.3. Questions about syntactic reconstruction 360
 - 4.4.4. The correspondence problem 363
 - 4.4.5. The 'pool of variants' problem 367
 - 4.4.6. Parametric comparison 368
 - 4.4.7. Conclusion 375
 - 4.5. Conclusion to Chapter 4 376
 - Further reading 377
 - 5. **Contact, creoles, and change** 383
 - Introduction 383
 - 5.1. Second-language acquisition, interlanguage, and syntactic change 384
 - 5.2. Contact and substrata 389
 - 5.2.1. Introduction 389
 - 5.2.2. Contact and word-order change in the history of English 391
 - 5.2.3. Substratum effects: Hiberno-English and Welsh English 399
 - 5.2.4. A 'borrowing scale' 404
 - 5.2.5. Conclusion 405
 - 5.3. Creoles and creolization 406
 - 5.3.1. Introduction: pidgins and creoles 406
 - 5.3.2. The Language Bioprogram Hypothesis 407
 - 5.3.3. The substratum/relexification hypothesis 419
 - 5.3.4. Conclusion: how 'exceptional' are creoles? 425

5.4. Language creation in Nicaragua 427

5.5. Conclusion to Chapter 5 438

Further reading 440

Epilogue 443

Glossary 445

References 457

Index of Subjects 491

Index of Names 501

Index of Languages 505

List of Tables

- 1.1 Synchronic verb-movement parameters 54
- 1.2 Parameters of Verb-movement in older Germanic, Romance, and Celtic languages 64
- 2.1 Verbal agreement inflection in Middle English 136
- 4.1 Occurrences of null and overt subjects in EP and PB 336
- 4.2 Parameter grid for nominal syntax 370

List of Figures

- 4.1 An idealized graphical change 297
- 4.2 Auxiliary *do*. Percentage of *do* forms in different types of sentence, 1500–1700 311
- 4.3 Social stratification of (r) in New York City 317
- 4.4 The rate of overt pronominal subjects in the nineteenth and twentieth centuries 338
- 4.5 Result of application of Kitsch to the data in Table 4.2 372

List of Boxes

- BOX 1.1: Technical aspects of movement 43
- BOX 1.2: Verb second in Old English 60
- BOX 1.3: The interpretation of *any* 73
- BOX 1.4: Cross-linguistic variation in negative concord 75
- BOX 1.5: Cross-linguistic variation in overt wh-movement 84
- BOX 1.6: The significance of wh-movement 88
- BOX 2.1: Merge and the LCA 190
- BOX 4.1: Ergative case marking 306

Preface

I'd like to thank the following people for their help, at different times and in different ways, with this book: Roberta d'Alessandro, Bob Berwick, Theresa Biberauer, Anna Cardinaletti, Lucia Cavalli, Chris Cummins, Teresa Guasti, Anders Holmberg, Nina Hyams, Judy Kegl, Ruth King, Adam Ledgeway, Glenda Newton, Ilza Ribeiro, Luigi Rizzi, Anna Roussou, Bonnie Schwartz, Christina Sevdali, Michelle Sheehan, Nigel Vincent, David Willis, the reviewers at Oxford University Press, and John Davey. All the errors are, as ever, entirely mine. I'd also like to thank the University of Cambridge for giving me sabbatical leave in 2005–6, which made it possible for me to finish the book.

*Downing College
Cambridge
January 2006*

List of Abbreviations and Acronyms

AAVE	African-American Vernacular English
ABS	absolute
ACC	accusative case
AN	adjective-noun
AP	adjective phrase
ASL	American Sign Language
BC	blocking category
BDT	Branching Direction Theory
BP	Brazilian Portuguese
CCH	Cross-Categorial Harmony
CL	clitic
CP	complementizer phrase
DAT	dative case
DP	determiner phrase
DR	Diachronic Reanalysis
EME	Early Middle English
ENE	Early Modern English
EP	European Portuguese
EPP	Extended Projection Principle
ERG	ergative
FC	free choice
HC	Haitian Creole
HPSG	Head-Driven Phrase Structure Grammar
IE	Indo-European
IP	inflection phrase
ISN	Idioma de Señas Nicaragüense
LAD	Language Acquisition Device
LCA	Linear Correspondence Axiom
LSN	Lenguaje de Señas Nicaragüense
ME	Middle English
MidF	Middle French
NA	noun-adjective

NE	Modern English
NOM	nominative case
NP	noun phrase
NPI	negative polarity item
NSP	Natural Serialisation Principle
OE	Old English
OF	Old French
ON	Old Norse
OV	object-verb
P&P	principles and parameters
PAC	Probably Approximately Correct
PEI	Prince Edward Island
PF	phonological form
PIE	Proto-Indo-European
PL	plural marker
PLD	primary linguistic data
PP	prepositional phrase
PRT	particle
PSN	Pidgin de Señas Nicaragüense
QI	quantity-insensitive
QP	quantifier phrase
QS	quantity-sensitive
SCL	subject clitic
SOV	subject-object-verb
SVO	subject-verb-object
TMA	tense/mood/aspect
TP	tense phrase
UG	Universal Grammar
VEPS	Very Early Parameter Setting
VO	verb-object
VP	verb phrase
VSO	verb-subject-object

*To all parameter-setting devices everywhere,
but most especially Julian and Lydia*

Introduction

Just under one thousand years ago, a monk named Ælfric translated the Latin Vulgate Bible into English. Here are a few lines from his translation of the passage in Genesis (3:1–4) describing the temptation of Eve:¹

Ēac swelce sēo nādre wæs gēappre þonne ealle þā ððre nīetenu þe God gēworhte ofer eorðan; and sēo nādre cwæð tō þām wīfe: ‘Hwȳ forbēad God ēow þæt gē neāten of ælcum trēowe binnan Paradīsum?’ Þæt wīf andwyrde: ‘Of þara trēowa wæstmte þe sind on Paradīsum wē etað: and of þæs trēowes wæstmte, þe is onmiddan neorxenawange, God bebēad ūs þæt wē ne āten, ne wē þæt trēow ne hrepoden þȳ læs wē swulten.’ Þā cwæð sēo nādre eft tō þām wīfe: ‘Ne bēo gē nāteswhōn dēade, þeah gē of þām trēowe eten.’

This is written Old English (OE), the language spoken in most of Anglo-Saxon England in various varieties – the standard one, illustrated above, known as West Saxon – from the time of the Anglo-Saxon invasions of the island of Britain in the fifth century until, according to the usual chronology, the Norman invasion in 1066. After this Norman French became the language of the ruling class; the English of the period 1066–1500 is conventionally known as Middle English (ME). Early Modern English (ENE) began in 1500, though it is sometimes dated from the introduction of printing into England in 1476, and Modern English (NE) in 1700. For speakers of Modern English, Old English appears to be a foreign language; to the untrained eye passages such as the above are indecipherable.

¹ The text is taken from Mitchell and Robinson (1992: 174). Details regarding the source of the text are given by Mitchell and Robinson (1992: 173). We have followed Mitchell and Robinson’s ‘normalization’ of the orthography and accents (see Mitchell and Robinson (1992: 11–12)). On Ælfric’s life and work, with particular emphasis on his authorship of the first grammar of Latin written in English, see Law (2003: 193–5).

This is, if you think about it, an odd state of affairs. We designate the language of the above passage as English, but recognize that no untrained speaker of today's English can understand it.² The reason for this is, of course, that languages change; and, as our passage illustrates, they can change almost out of recognition in the course of a millennium. This book aims to present some recent ideas regarding certain aspects of this phenomenon of language change, in the context of an influential general theory of language.

The particular aspect of language change that this book is concerned with is syntactic change, change in the ways in which words and phrases are combined to form grammatical sentences. If we update all the other aspects (vocabulary, orthography, etc.) of our passage from Ælfric above, but keep the syntax the same as the Old English, we have something like the following:

Also such the snake was deceitfuller than all the other beasts that God made on earth; and the snake said to the woman: 'Why forbade God you that ye not eat of each tree in Paradise?' The woman answered: 'Of the trees' fruit that are in Paradise we eat: and of the tree's fruit, that is in-the-middle-of Paradise, God bade us that we not eat, nor that we the tree not touch lest we die.' Then said the serpent back to the woman: 'Not be ye not-at-all dead, though that ye of the tree eat.'

This is a word-for-word rendering of the passage into NE (hence 'in-the-middle-of' is hyphenated, as it corresponds to the single OE word *onmid-dan*, and similarly 'not-at-all' for *nāteswhōn*). Although it is now comprehensible, my rendering brings to light a number of syntactic differences between Ælfric's English and today's. Of these we can note the form of the question the serpent puts to Eve: '*Why forbade God you . . . ?*' In NE, main

² Here is the King James Bible version of the same passage. This variety of English is of course somewhat archaic, representing a literary variety of the early seventeenth century; it is nonetheless relatively comprehensible for modern readers:

Now the serpent was more subtil than any beast of the field which the LORD God had made. And he said unto the woman, Yea, hath God said, Ye shall not eat of every tree of the garden?

2 And the woman said unto the serpent, We may eat of the fruit of the trees of the garden:

3 But of the fruit of the tree which *is* in the midst of the garden, God hath said, Ye shall not eat of it, nor shall ye touch it, lest ye die.

4 And the serpent said unto the woman, Ye shall not surely die.

(from Mitchell and Robinson (1992: 175))

verbs like *forbid* do not invert with subjects in questions, the auxiliary *do* being used instead (i.e. ‘*Why did God forbid you ...?*’). We see another instance of ‘main-verb inversion’ – but this time not in a question – in the penultimate line: ‘*Then said the serpent back to the woman*’. Another striking difference occurs twice in the last two sentences: Eve says ‘... *that we the tree not touch*’ and the serpent, in his reply, says ‘*though that ye of the tree eat*’. Here we see the order subject (*we/ye*), object ((*of*) *the tree*), verb (*touch/eat*); this is an order which NE does not usually allow but which is usual in OE subordinate clauses. Further differences can be observed. Look, for example, at the position of the negative word *ne*: I have translated this as ‘not’; although in fact *ne* died out in late ME, and NE *not* derives from OE *nan wuht* ‘no wight’ (‘no creature’). Notice too the occurrence of ‘that’ in various places where it is not allowed in NE, such as following ‘though’ in the last sentence. I have not attempted to represent the OE case marking on nouns and articles, but this can be observed in the different forms of the word I have translated as ‘the’: *sēo* (nominative singular masculine), *þæt* (nominative/accusative singular neuter), *þāra* (genitive plural), *þæs* (genitive singular masculine/neuter), *þām* (dative singular masculine/neuter), etc. These case markings have all but disappeared in NE, a development which, although in itself a morphological or phonological change, may have affected English syntax.

So we can see that English syntax has changed in a number of ways in the past thousand years. But during that period the language has been passed on from generation to generation in the normal way, first in England and later in the various countries where English speakers settled. Children have learnt the language at their mothers’ knees, and there is no good reason to think that invaders or other foreign influences caused the kinds of changes we have just observed, with the possible exception of the Norse invaders of the ninth to eleventh centuries (see §5.2.2). In particular, although English has absorbed a great deal of vocabulary from French and Latin in the past millennium, there is no evidence that either of these two languages has influenced English as far as the types of changes we have just observed are concerned. So how and why did these changes take place? That is the central question this book will address.

English is by no means untypical as far as syntactic change is concerned. The example of Ælfric’s Bible translation could easily be replicated by comparing an excerpt from a twelfth-century *chanson de geste* with Modern French, or by comparing Plato’s syntax with Modern Greek, or the Vedic

hymns with Modern Hindi. Like all other types of language change, syntactic change can be observed wherever we compare surviving ancient texts with those in a corresponding modern language. As has often been observed, change appears to be almost an inherent feature of all aspects of language. Language, to use McWhorter's (2001: 52) phrase, appears to show a kind of 'structured variation'. The purpose of this book is to present some recent ideas concerning this structured variation in syntax and apply them to change over time. To do this, we must develop a general theory of the nature of the structures and of the nature of the variations.

The theory of syntactic variation is the object of the first chapter, and so I will say no more about it here. Concerning the nature of syntactic structure, I will adopt what is arguably the most influential theory of recent years: that developed by Noam Chomsky and his associates and usually known as **generative grammar**.³ The most recent variant of generative grammar is known as the **Minimalist Program**, and I will assume a version of this in what follows. However, since my goal here is neither to develop nor to defend this particular version of generative syntax, I will try to keep the technical details to a minimum. I hope that those who are fully conversant with these details will not see my approach as too simplistic, and that those who are unfamiliar with them will not be deterred.

Two aspects of Chomsky's thinking about language are central to what follows, and we must be explicit about these. The first is the idea that sentences can be exhaustively divided up into smaller constituents, down at least to the level of the word,⁴ and that the basic combinatorial principles are **discrete**, **algorithmic**, **recursive** and purely **formal**. By 'discrete' I mean that the elements of syntax are clearly distinguished from one another: clines, squishes, fuzzy sets, and continua play no role. By 'algorithmic', I mean that syntactic structures can be determined in an explicit, step-by-step fashion. By 'recursive', I mean that syntactic operations can apply to their own output, thus in principle creating infinite structures from a finite set of symbols and types of operations. And by 'formal' I mean that syntactic operations are not directly determined by semantics, but can be

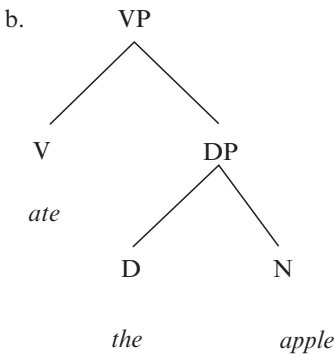
³ Boldfaced items in the text are defined in the Glossary.

⁴ This idea is not original to Chomsky. It was an aspect of the American structuralist school of linguistics which was dominant in the United States prior to the 1950s, and has older historical antecedents; Seuren (1998: 219) traces it back to Wundt (1880).

seen as operations which manipulate symbols independently of any denotation those symbols may have.

The simplest way to illustrate these ideas is in terms of the basic operation **Merge**, proposed in Chomsky (1995). Merge combines two syntactic elements (in the simplest case, two words) into a more complex entity which consists of those two elements and its label, the label being determined by one of the two elements. For example, Merge may combine the noun *apple* with the determiner *the*, forming the larger phrase *the apple*. The resulting phrase is usually regarded as a determiner phrase (DP) in current work, reflecting the assumption that the label of the larger unit formed by Merge is contributed by the determiner *the*. Merge may then combine the verb *ate* with the DP *the apple*, giving the phrase *ate the apple*, which is taken to be a verb phrase (VP) – the verb determines the label of the larger unit in this case. The structure that results from these operations can be represented as a **labelled bracketing**, as in (1a), or as a **tree diagram**, as in (1b):

(1) a. [VP [V ate] [DP [D the] [N apple]]]



These two representations are entirely equivalent, the choice of either one being determined by didactic or typographical considerations. Merge is discrete, in that it combines distinct elements (words, categories); algorithmic, in that it can be seen to apply in a mechanical, step-by-step fashion (and it can be formalized rather more precisely than I have done here – see Chomsky (1995: 241ff.)); recursive, in that it applies to its own output, as our example illustrates with the DP formed by Merge being itself part of the input to the next operation of Merge forming the VP; and formal in that reference is not made to the meaning of the symbols combined.

The second aspect of Chomsky's thinking which is important for our purposes has two components: that the fundamental principles of syntax are universal, and that they may therefore reflect some aspect of human cognition. These two points are logically distinct, although they naturally go together. The first idea is that operations like Merge are not specific to any particular language but are **formal universals of language**. This is a radical and thought-provoking idea,⁵ which has given rise to much debate over many years. It implies that principles such as Merge must have been operational in Ælfric's English, Plato's Greek etc. every bit as much as they are in present-day English. Assuming formal universals in this way means that our approach to historical questions adheres to the **uniformitarian hypothesis**, the idea that 'the languages of the past . . . are not different in nature from those of the present' (Croft 2003: 233). Rather than attempt to justify Chomsky's radical idea here, I hope that the chapters to follow will show that this idea has a number of very interesting empirical and conceptual consequences.

Chomsky's further proposal that the formal universals of language represent an aspect of human cognition has given rise to even more controversy. What is most relevant in the present context is that it allows us to relate syntactic structure to children's acquisition of their first language. During the early years of life these universals are put into action as the child develops the capacity to speak and understand. There are two ways to think about how this may happen. On the one hand, if the universals themselves must be acquired, then this of course must happen during language development. On the other hand, if the universals are inherited (since language is common to all – and only – humans, and inherent universals may be thought of as part of the human genome along with other specifically human features), then they are simply applied to the task of language acquisition. The celebrated argument from the **poverty of the stimulus** (which I will briefly review at the beginning of Chapter 1) asserts that the second of these views is the more plausible of the two. But, whether

⁵ Again, I do not mean to imply that it is new. The concept of universal grammar was discussed by the Cartesian Port-Royal grammarians in the seventeenth century, and was arguably implicit in the thinking of the medieval speculative grammarians (the *modistae*) (Law 2003: 264). Chomsky (1965, Chapter 1; 1966) discusses his own view of some of the historical antecedents of his ideas on this and other matters. Chomsky (1966) is critically reviewed by Aarsleff (1970) and is also commented on by Simone (1998), among many others (see the references given in Simone (1997: 150)).

we accept this or not, it is clear that *differences* between languages must be acquired as part of the process of first-language acquisition. Going back to our notion of structured variation, we see that, while the universal structures may be either inherited or acquired, the variation must be acquired. Since historical change is variation in time, this in turn implies a connection between historical change and language acquisition. This is an old idea (Paul 1920); see Harris and Campbell (1995, 31); Morpurgo-Davies (1998: 248–51), which has been taken up most notably and influentially by Lightfoot (1979; 1991; 1999) in the context of generative grammar. In much of what follows we will explore the ramifications of this idea.

To sum up, two of Chomsky's ideas are central to the discussions to follow: the idea that there are formal universals of syntax and the idea that these universals are an aspect of human cognition. I have sketched the two ideas here, but I have not attempted to do them justice. Recent introductions to the theory of syntax, all of which go over these points to a greater or lesser extent, are Adger (2003); Carnie (2000); Haegeman (1994, 2005); Hornstein, Nunes, and Grohmann (2005); Lasnik, Uriagereka, and Boeckx (2004); Ouhalla (1994); Radford (2004); and Roberts (1996). A basic grounding in Chomsky's ideas about the cognitive status of language is presented, for example, in Cook and Newson (1996) or N. Smith (2004), while Chomsky (2000, 2002, 2005a) goes into these questions in more detail and, in the case of the latter two, with specific reference to the Minimalist Program. However, this book is not intended to build directly on the textbooks in the sense of providing more sophisticated analytical techniques or theory-internal reflections. Discussions of the relative merits of the Minimalist Program, or of any other designated approach to the nature of the formal universals of syntax, will not figure: I will simply adopt an informal version of minimalism.

In the widest sense, then, the goal of this book is to illustrate how Chomsky's two ideas just summarized can form the basis for the study of historical syntax. These ideas can shed light on how and why English has changed since Ælfric's time in the ways we observed above, and allow us to integrate our account of these changes with a general theory of structure and variation in syntax. I hope that this book will provide a clear conception of the implications of Chomskyan thinking for traditional questions in historical linguistics and a different perspective on the nature of Universal Grammar and first-language acquisition. For those already familiar with Chomskyan syntax, I hope it will provide an illustration of the importance and relevance of syntactic change for our conception of how grammatical

systems vary syntactically over time and how such systems are acquired. In this sense, the book is written from an explicitly Chomskyan perspective, although the emphasis is on the interpretation and extension of Chomsky's thinking, rather than on the defence, exegesis, or criticism of specific proposals – technical or philosophical – in Chomsky's writings.

Finally, I should point out what this book is *not* intended to do. It is not intended as a manual for syntactic analysis; the textbooks cited above fulfil this function. Neither is it intended as a guide for doing historical work, whether of a traditional philological kind or of a computational, corpus-driven kind. Instead, as stated above, the book is intended as an introduction to a particular area of linguistic theory.

Further reading

At the end of each chapter, I will give a few details and comments on the more important works mentioned. Naturally, a number of works are mentioned in more than one chapter; I will comment on each work at the end of the first chapter in which it is mentioned. Thus, if the reader does not find a comment on a work at the end of a later chapter, the preceding chapters should be checked. Not every single reference mentioned in the text is commented on in these sections, but all of the more significant and useful works are.

The further reading mentioned in this chapter falls into various categories:

Works on the history of linguistics

Law (2003) is a recent and very thorough overview of the history of linguistic thought in Western Europe from antiquity to 1600. **Chomsky (1966)** contains Chomsky's own assessment of the seventeenth-century antecedents to his thinking on the nature of language, mind, and grammar. **Aarsleff (1970)** is a very critical assessment. **Morpurgo-Davies (1998)** surveys the history of linguistics in nineteenth-century Western Europe, and provides a valuable perspective on the development of modern historical linguistics. **Seuren (1998)** is a very interesting history of western linguistics, usefully combined with a history of logic in Western Europe. The views expressed on generative grammar are somewhat idiosyncratic, however.

Textbooks on syntax

Adger (2003) is an introduction to minimalist syntax, which presupposes no prior knowledge of earlier versions of syntactic theory. **Radford (2004)** is a very comprehensive introduction, again presupposing no prior knowledge of syntax. **Carnie (2000)** is a more general introduction and combines elements of minimalist syntax with those of the earlier government-and-binding theory, as do **Ouhalla (1994)** and **Roberts (1996)**. **Haegeman (1994)** is the most comprehensive introduction to government-and-binding theory available, and **Haegeman (2005)** is a general introduction to syntactic theory. **Hornstein, Nunes, and Grohmann (2005)** and **Lasnik, Uriagereka, and Boeckx (2004)** are up-to-date introductions to the technical aspects of the Minimalist Program written by teams of leading experts in the field.

Chomsky's work and introductions to it

Chomsky (1965) remains in many ways the foundational text of generative grammar; Chapter 1 of this book is arguably Chomsky's fullest and most lucid introduction to the goals of generative grammar to date. **Chomsky (1995)** is a collection of papers from the early 1990s, including (Chapter 3) the first exposition of the Minimalist Program, and (Chapter 4) some very important refinements of those initial ideas. The technical notions of minimalism are further developed and refined in **Chomsky (2000)** and elsewhere (see the further reading in later chapters), while **Chomsky (2002; 2005a)** present the conceptual background to the Minimalist Program. **Cook and Newson (1996)** is an accessible introduction to Chomsky's thinking on Universal Grammar, although some of the ideas presented are a little outdated. **N. Smith (2004)** is more up to date, and covers Chomsky's thinking on a range of issues, including politics.

Historical linguistics

Lightfoot (1979) is arguably the foundational text in diachronic generative syntax, and the direct inspiration for much of the material in this book. **Lightfoot (1991)** develops a number of the central ideas of the earlier work,

as well as introducing the notion of ‘degree-0 learnability’, which will play a role in our discussion, notably in Chapter 3. **Lightfoot (1999)** restates and elaborates a number of the ideas from the earlier works. **Harris and Campbell (1995)** is a very interesting survey of the issues in diachronic syntax from a non-Chomskyan theoretical perspective, and contains a number of clarifications of core questions, as well as some interesting novel proposals. **Mitchell and Robinson (1992)** is the most comprehensive introduction to Old English and Anglo-Saxon literature and culture available. **Paul (1920)** is a classic statement of the concepts and methods of historical linguistics, written by a major neogrammarian. This work remains influential to this day.

1

Comparative and historical syntax in the principles-and parameters-approach

Introduction	11	1.4. Negative concord	64
1.1. UG and variation in grammatical systems	19	1.5. Wh-movement	81
1.2. The null-subject parameter	24	1.6. Head-complement order	92
1.3. Verb-movement parameters	40	1.7. Summary	108
		Further reading	110

Introduction

In this chapter I will present the way in which syntactic variation is analysed in current theory. The central notion is that of **parameter of Universal Grammar**, a term which is fully explicated in §1. The rest of the chapter is devoted to illustrating certain parameters of Universal Grammar, with examples taken from both the synchronic and the diachronic domains. We establish that the analytic device which has been used to describe synchronic variation across languages can also be used to describe diachronic changes between different stages of the same language.

Before introducing parametric variation, however, we need to be more precise about what does *not* vary, i.e. about the nature of the formal universals of syntax that were mentioned in the Introduction.

Chomsky has always argued that one of the goals of linguistic theory is to develop a general theory of linguistic structure that goes beyond simply describing the structures of individual languages (see Chomsky (1957: 50); Matthews (2001: 100ff.) and references given there). In other words, a major concern of linguistic theory is to develop a characterization of a possible human grammar. To do this, we elaborate a theory of the formal universals of human language, known as **Universal Grammar**, or UG. This is taken to be the set of grammatical principles which both make human language possible and define a possible human language. Thus, UG embodies the essential invariant parts of the structure of language. Whilst our main concern in this book is with syntax, UG also contains principles related to phonology, morphology, and semantics. Whether these aspects of language are subject to parametric variation in the same way as syntax is an open question; there is some reason to think that this is true of phonology and morphology (see the discussion of Dresher (1999) in §3.3, for an example of a phonological parameter), while semantics may not be subject to variation. However, owing to my own lack of relevant expertise, I will leave these other subsystems of UG aside and concentrate on syntax.

One important way in which syntax makes human language possible has to do with its recursive nature. As mentioned in the Introduction, recursion makes it possible to construct infinite structures from a finite number of elements. The recursive nature of syntax is a necessary component of what Chomsky has called the ‘creative aspect of language use’: the fact that humans are able to produce and understand utterances that have never been produced before. This formal property of natural-language syntax allows us to give expression to our freedom of will.

In saying that UG defines a possible human language, I mean that UG is intended as a general theory of the structure of human language, and not simply an account of the structure of the set of languages that happens to exist at this – or any other – historical moment. To be more precise, UG is intended to give an account of the nature of human *grammar*, rather than *language*; the notion of grammar is more precise and less subject to confusion due to social, political, and cultural factors than that of language. Moreover, whilst a language can be thought of just as a set of strings of symbols, a grammar is more abstract, being the device which determines

which sets of symbols are admitted in the language. In other words, even if we had at our disposal the means, both intellectual and practical, to write an exhaustive description of the grammar of every language currently spoken and every language for which textual evidence survives, any resulting inductive distillation of the results of such a survey would not yield UG. It could yield an extensional definition of the common features of all currently (and recently) existing grammars, and would be universal in this weak sense. But what UG aims for is an intensional characterization of the class of human grammars: a characterization of what makes a grammar what it is. UG should tell us what the defining properties of any possible human grammar are.

A natural question to ask is whether UG is a purely abstract entity (for example, a set of some kind) or whether it has some physical or mental existence. Chomsky's view has always been that UG has mental reality, in that it corresponds to an aspect of the mind, the **language faculty**. We can define UG as our theory of the language faculty, the mental faculty or faculties which both facilitate and delimit the nature of grammar. This view has the advantage that UG can now be seen as being in principle a theory of an aspect of physical reality; the language faculty – as a mental reality – is physically instantiated in the brain (somehow – a number of complex philosophical, psychological, and neurological issues arise here).¹ Furthermore,

¹ Recently, Hauser, Chomsky, and Fitch (2002) have suggested, following the leading ideas of the Minimalist Program (Chomsky 2004; 2005a, b), a distinction between on the one hand the language faculty in the broad sense, i.e. syntax along with phonology and semantics, and on the other hand the language faculty in the narrow sense, i.e. just syntax. They suggest that much of the language faculty in the broad sense 'is based on mechanisms shared with non-human animals', while syntax, above all because of its digital, recursive character (as discussed in the Introduction), may be the crucial language-specific component (Hauser *et al.* 2002: 1573). In fact, they entertain the possibility that even recursion may have an evolutionary origin outside language (1578). Furthermore, a central idea behind the Minimalist Program is that all aspects of the language faculty may be shaped by features of optimal design which are language-independent (see Chomsky (2005a: 9ff.; 2005b: 1–4)). For these reasons, it may not be correct to think in terms of a specialised 'mental module' for language, although there is some evidence from language pathology for this (see in particular N. Smith (2004) and N. Smith and Tsimpli (1995)). There is also evidence for a critical period specific to language acquisition, as we shall see in §5.4, which may in turn favour of the postulation of a 'language module', although not as a logical necessity. Clearly, though, the claim that language is a facet of cognition and

there is nothing mysterious about the idea that the language faculty may be genetically inherited – the **innateness hypothesis**. This is the idea that the particular aspects of cognition which constitute the language faculty are a consequence of genetic inheritance, and it is no more or less surprising and problematic than the general idea that cognition is to some degree genetically facilitated. And if cognition is physically instantiated in the brain (somehow), then the claim is just that aspects of the physical functioning of the brain are genetically inherited.

To recapitulate: it is a goal of linguistic theory to attempt to develop a general characterization of a possible human grammar. It is reasonable, although not a matter of logical necessity, to take this characterization to be a reflection of some aspect of how the mind works, i.e. as a facet of human cognition which we call the language faculty. If cognition has a physical basis in the brain, then so does the language faculty. Finally, it may be that the language faculty is genetically inherited; that some aspect of the human genome determines its existence in all normal humans.

The innateness hypothesis is highly controversial. As mentioned in the Introduction, the principal argument for it is the poverty-of-the-stimulus argument. Here I will briefly summarize this argument (for a more detailed presentation, see Roberts (1996: 265–71); N. Smith (2004: 38ff.); Jackendoff (2002: 82–7); and, in particular, Guasti (2002: 5–18); Pullum and Scholz (2002) present a very strong version of the poverty-of-the-stimulus argument, which they subject to a detailed critique). As its name implies, the poverty-of-the-stimulus argument is based on the observation that there is a significant gap between what seems to be the experience facilitating first-language acquisition (the **primary linguistic data**, PLD henceforth) and the nature of the linguistic knowledge which results from first-language acquisition, i.e. one’s knowledge of one’s native language. The following quotation summarizes the essence of the argument:

The astronomical variety of sentences any natural language user can produce and understand has an important implication for language acquisition ... A child is exposed to only a small proportion of the possible sentences in its language, thus limiting its database for constructing a more general version of that language in its own mind/brain. This point has logical implications for any system that attempts to

physically instantiated in the brain does not entail the postulation of a language module. For a response to Hauser, Chomsky, and Fitch (2002), defending the conception of modularity, see Pinker and Jackendoff (2005); Jackendoff and Pinker (2005).

acquire a natural language on the basis of limited data. It is immediately obvious that given a finite array of data, there are infinitely many theories consistent with it but inconsistent with one another. In the present case, there are in principle infinitely many target systems . . . consistent with the data of experience, and unless the search space and acquisition mechanisms are constrained, selection among them is impossible . . . No known ‘general learning mechanism’ can acquire a natural language solely on the basis of positive or negative evidence, and the prospects for finding any such domain-independent device seem rather dim. The difficulty of this problem leads to the hypothesis that whatever system is responsible must be biased or constrained in certain ways. Such constraints have historically been termed ‘innate dispositions,’ with those underlying language referred to as ‘universal grammar.’

(Hauser, Chomsky, and Fitch 2002: 1576–7)

Similarly, in introducing the general question of the nature of the learning problem for natural languages, Niyogi (2004: 16) points out that the basic problem is

the inherent difficulty of inferring an unknown target from finite resources and in all such investigations, one concludes that *tabula rasa* learning is not possible. Thus children do not entertain every possible hypothesis that is consistent with the data they receive but only a limited class of hypotheses. This class of grammatical hypotheses *H* is the class of possible grammars children can conceive and therefore constrains the range of possible languages that humans can invent and speak. It is Universal Grammar in the terminology of generative linguistics.

As an illustration of the complexity of the task of language acquisition, consider the following sentences:

- (1) a. The clowns expect (everyone) to amuse them.
- b. The clowns expected (everyone) to amuse themselves.

If *everyone* is omitted, the pronoun *them* cannot correspond to *the clowns*, while if *everyone* is included, this is possible. If we simply change *them* to the reflexive pronoun *themselves*, as in (1b), exactly the reverse results. In (1b), if *everyone* is included, the pronoun *themselves* must correspond to it. If *everyone* is left out, *themselves* must correspond to *the clowns*. (One might object that facts such as these are semantic, but they are usually considered to be partially determined by syntax – the usual analyses of these phenomena are described in the textbooks cited in the Introduction). The point here is not how these facts are to be analysed, but rather the precision and the subtlety of the grammatical knowledge at the native speaker’s disposal. It is legitimate to ask where such knowledge comes from.

Another striking case involves the interpretation of missing material, as in (2):

- (2) John will go to the party, and Bill will – too.

Here there is a notional gap following *will*, which we interpret as *go to the party*; this is a ‘missing’ VP, and the phenomenon is known as VP-ellipsis. In (3), we have another example of VP-ellipsis:

- (3) John said he would come to the party, and Bill said he would – too.

Here there is a further complication, as the pronoun *he* can, out of context, correspond to either *John* or *Bill* (or an unspecified third party). Now consider (4):

- (4) John loves his mother, and Bill does – too.

Here the gap is interpreted as *loves his mother*. What is interesting is that the missing pronoun (the occurrence of *his* that isn’t there following *does*) has exactly the three-way ambiguity of *he* in (3): it may correspond to *John*, to *Bill* or to a third party. Example (4) shows we have the capacity to apprehend the ambiguity of a pronoun which is not pronounced.

The above cases are examples of native grammatical knowledge. The basic point in each case is that native speakers of a language constantly hear and produce novel sentences in that language, and yet are able to make very subtle judgements of interpretation and ambiguity. They are also able to distinguish well-formed sentences from ill-formed ones. Here is a further example, uttered while planning a party, for example. This example is based on Radford (2004: 15):

- (5) Who did he think was likely to drink what?

This sentence has a natural interpretation, known as the ‘pair-list’ interpretation, according to which an answer to ‘who’ and an answer to ‘what’ are paired (i.e. ‘He thought John was likely to drink vodka, Mary gin, Bill orange juice,’ etc.). We understand the sentence this way naturally, and moreover we immediately understand that *he* must be disjoint from *who*. Also, we can recognize the following variants of this example as ungrammatical (indicated by an asterisk), even if they are trying to mean the same thing:

- (5') a. *Who did he think that was likely to drink what?
b. *What did he think who was likely to drink?
(cf. What did he think John was likely to drink?)

- c. *Who was he thought likely to drink what?
- d. *Did he think who was likely to drink what?

The question is why and how we are able to distinguish previously unheard examples like (5) from ungrammatical but extremely similar ones such as (5'). In first-language acquisition, **negative evidence** – information about ungrammatical sentences – is unavailable; children may be exposed to ungrammatical sentences but they are not told that they are ungrammatical; where explicit instruction is intended, it appears to be either ignored or misunderstood. Meaning probably isn't much help in distinguishing the examples in (5), as the sentences in (5') mean the same as those in (5), to the extent they mean anything, which (5'a) pretty clearly does. This knowledge must either come from experience or from within. If we truly have no experience of novel sentences like (5), then it must come from within. Moreover, if the knowledge of these properties of English comes from within, it must represent some aspect of UG, as there is no genetic disposition to English. Here we see the links between the poverty of the stimulus, the postulation of an innate language faculty, and UG.

To put it another way, if we deny that knowledge of grammar of the type illustrated in (1)–(5) can be innate, then we must maintain that the conditions of language acquisition and the nature of our minds (minds by hypothesis lacking any special predisposition to grammatical knowledge) are such that we are able to glean subtle aspects of the interpretation of pronouns purely from experience, including absent pronouns as in (4); we must also be able to distinguish sentences like (5) from non-sentences like (5'). Despite much criticism of the poverty-of-the-stimulus argument (see Pullum and Scholz (2002) and the references given there), no clear account of why and how native speakers can do any of this has emerged. On the other hand, introductory textbooks of the kind referred to earlier offer such an account in terms of an innate UG.

Of course, a natural response is to say that, while we may never have heard (5), we have heard plenty of examples like it. But here we must be very clear about what 'like (5)' actually means. If 'like (5)' means 'containing the same, or nearly the same words, as (5)' then of course (5') are very like (5); these examples contain exactly the same words as (5) in all cases except one. But the examples in (5') are ungrammatical while (5) is grammatical. Construing 'like (5)' in any other sense involves attributing knowledge of some aspect of syntactic structure to speakers who recognize the difference

between (5) and (5'), and this is exactly what the poverty-of-the-stimulus argument is trying to explain. Thus the question of the mental status and the origins of that knowledge is begged.

The idea of some kind of superficial resemblance among sentences as informing language acquisition has underlain many behaviourist theories of acquisition. Chomsky (1959) showed how one rather well-worked-out behaviourist theory of language acquisition was doomed to failure. More recently, Guasti (2002: 10–17) provides a detailed discussion of why mechanisms such as imitation, reinforcement, and association are unable to account for the first-language acquisition of such aspects of grammar. Moreover, there is evidence that first-language acquisition takes place on the basis of 'positive evidence' only, in the sense that children do not have access to information regarding what is *not* allowed; they only hear examples of what *is* possible (see Guasti (2002: 3–4) and the references given there; this issue is complex as it involves making assumptions regarding what children 'do' with what they hear, about which almost nothing is known). Also, language acquisition takes place in a largely uniform way across children from different social groups and language backgrounds, does not rely on explicit instruction and happens very quickly given the complexity of the task and the relatively rudimentary nature of more general reasoning and other cognitive skills at an early age. Most of first-language acquisition is effectively accomplished by the age of six.

The poverty-of-the-stimulus argument asserts that, given the factors mentioned above, it is highly implausible to think that there is no predisposition to language at all. If there is a 'predisposition to language', then some aspect of linguistic knowledge is innate. In the absence of any account of how grammatical knowledge like that illustrated in (1–5) – and a myriad of similar examples (see Anderson and Lightfoot (2002: 198–206); Crain and Pietroski (2002); Fodor and Crowther (2002); Jackendoff (2002: 82–7), and the references given in these sources) – may be determined purely on the basis of experience by a mind with no predisposition to language, we conclude that knowledge of language arises from the interaction of innate knowledge with relevant experience. This does not mean that UG directly determines facts of the type in (1)–(5) regarding ellipsis, anaphora, etc., but rather that such facts can be seen as consequences of fairly abstract innate principles interacting with experience. The question of the balance between innate knowledge and experience is difficult and complex; it is also to a considerable extent an empirical matter, i.e. it cannot be determined purely

by theoretical speculation. (This point is made by Pullum and Scholz (2002).) We will come back to this in §1.1 below. The important point is that the innateness hypothesis can provide a solution to the poverty-of-the-stimulus problem. As Guasti says:

The hypothesis that the language capacity is innate and richly structured explains why language acquisition is possible, despite all limitations and variations in learning conditions. It also explains the similarities in the time course and content of language acquisition. How could the process of language acquisition proceed in virtually the same ways across modalities and across languages, if it were not under the control of an innate capacity?

(Guasti 2002:17)

For many years, Chomsky has argued for an innate language faculty, and takes UG to be the theory of this faculty. Here I will follow this view, in part because of the force of the poverty-of-the-stimulus argument as just given. In the chapters to follow, I will try to show that this point of view can be revealing for our understanding of language change.

1.1. UG and variation in grammatical systems

In the previous section we saw the reasons for postulating the existence of an innate language faculty. Poverty-of-stimulus considerations of the type outlined there lead us to think that the language faculty is richly specified. In fact, once we begin to think in terms of an innate language faculty, it becomes apparent that the more aspects of individual grammars we can ascribe to UG, the more we can simplify the account of language acquisition. Given that first-language acquisition requires explanation, for the reasons we saw in outline above, this is a good result.

Nevertheless, it seems that we cannot escape the fact that different languages have different grammars. We can easily observe that a sentence which is syntactically well-formed in one language may be ill-formed in some other language. Compare the following very simple sentences and non-sentences in English and German:

- (6) a. Tomorrow John will visit Mary.
b. *Morgen Johann wird besuchen Maria.
- (7) a. Morgen wird Johann Maria besuchen.
b. *Tomorrow will John Mary visit.

Example (6a) is a quite unremarkable English sentence, but its exact syntactic counterpart in German – a word-for-word translation with the words retaining their English order – is ungrammatical in German. Conversely, (7a) is a correct German rendering of the English (6a), but if we translate it back into English retaining the German word order, we arrive at the impossible (7b). The conclusion is clear: English syntax differs from German syntax. How are we to reconcile this conclusion with the postulation of a rich, innate UG?

One simple way to answer the question would be to say that English speakers and German speakers are genetically distinct: one aspect of this genetic difference is a difference in the respective language faculties, which has the consequence that English and German have different syntax. This gives rise to the differences observed in (6) and (7). However, this view cannot be maintained, since we have ample evidence from immigrant communities the world over that children of speakers of one language are perfectly able to become native speakers of the language of their adopted community. In the case of English and German, it suffices to point to the large numbers of German-speaking immigrants to the United States in the late nineteenth and early twentieth centuries whose descendants have, by now for several generations, been native speakers of English. This simple fact is incompatible with the idea that the syntactic differences observed in (6) and (7) are attributable to some genetic difference between English speakers and German speakers. (Of course, this view would also lead us to postulate genetic differences in cognition between different nationalities and ethnic groups, a highly dubious move on ethical grounds. As we see, however, there is no support for this position, and good evidence against it, in this as in many other domains).

If differences between the grammars of different languages cannot be accounted for in directly genetic terms, then how are we to account for them, given the assumption of an innate UG? It would seem that these differences are not part of the innate language faculty. However, it does not take much technical knowledge of syntax to be able to tell that the differences in word order between English and German in (6) and (7) involve fairly central aspects of syntax. (They involve at least the position of the verb and the position of the direct object, as we shall see in §1.3 §1.5.) Moreover, if we postulate that syntactic differences among languages are not part of the genetic endowment then we may expect these differences to be quite random in relation to UG. However, there are good reasons

to think that syntactic variation across languages is not random in this sense; this conclusion has been established by language typologists quite independently of the assumption of UG in the sense described in the previous section. For example, in (6) and (7) we can see that in English the verb always precedes the direct object; one aspect of the ungrammaticality of (7b) is the fact that the direct object precedes the verb (**Mary visit*). For this reason, English is referred to as a **VO (verb-object) language**. On the other hand, in German non-finite verbs generally follow their objects (cf. *Maria besuchen* in (7a) and the ungrammaticality of the reverse order in (6b)); German may therefore be considered a kind of **OV language**. Language typologists have shown that a number of other variant traits are correlated with VO vs. OV order (see Comrie (1989); Croft (2003); Song (2001) for introductions to **language typology**; we will return to these questions in more detail in §1.6, §2.5, and §3.5). It is thus now generally accepted that syntactic variation among languages is not random. For this reason, coupled with the fact that fairly central properties seem to vary, we want to ‘build variation in’ to our theory of UG.

What is required is a way of expressing syntactic variation within UG itself. This is achieved by adopting the notion of parameters of variation. The central idea is quite simple: associated with the invariant principles of UG there may be certain limited options which remain open, to be ‘filled in’, as it were, by experience. These options determine the parameters along which grammars may vary, and are thus known as the parameters of UG. In this view, UG consists of invariant principles and associated parameters. It is important to see that both the principles and parameters are innate, as is the range of options specified by the parameters. Experience is necessary only to fix the values of the parameters. The importance of this conception of the interaction of structure and variation is such that this approach to syntax as a whole has become known as principles-and-parameters theory (P&P theory).

We can illustrate the interaction of principles and parameters in an informal way using our examples of word-order differences between English and German seen in (6) and (7) above. We saw that English is a VO language and that German (at least where V is non-finite, a complication we leave aside in this illustration) is OV. In terms of P&P theory, we could say that UG principles determine the nature of V (the universal theory of syntactic categories, i.e. Nouns, Verbs, etc., would do this), the nature of O (the universal theory of **grammatical functions** such as subject, direct

object, indirect object, etc., would do this) and how they combine to form a VP (this may be determined in part by the nature of Merge, as we saw in the Introduction, and in part by the theory of grammatical functions, which would state for example that a determiner phrase (DP) merged with V is a direct object). So universal principles dictate that a Verb and its direct object combine by Merge to form a VP. The parametric option concerns the linear order of V and O – universal principles say nothing about this. Hence grammars may choose either of the two logical options: OV or VO. As we saw, German takes the former option while English takes the latter.

There are several points to note regarding this brief and somewhat simplified illustration. First, we see that the role of experience lies simply in determining the linear order of rather salient elements: verbs and their direct objects. The actual learning task is thus rather simple, and, impoverished though it may be, the stimulus is presumably not so defective that this information cannot be detected by language acquirers. So we reconcile poverty-of-the-stimulus considerations with cross-linguistic variation. This, in essence, is the great attraction of the P&P approach. It also specifies quite clearly the relation between experience and the innate faculty.

Second, a choice has to be made: not deciding is not an option. At the relevant level of abstraction, the task of acquisition of syntax consists purely in fixing the values of parameters in this way. The ability to acquire a given language may thus be construed as the ability to set the parameters of UG to determinate values. Each grammar must choose a value for each parameter, although certain values may be determined by default – I return to this point in §3.5. An implication of this is that all languages can and must be defined as OV or VO; ‘free word-order’ languages cannot exist, for example. It is a matter of debate whether this is in fact the correct view, but we see how the logic of P&P reasoning can lead us to this conclusion.

Third, options may be determined by ‘gaps’ in UG principles. This appears to be the case in our example of OV vs. VO: everything about the Merge of V and its object to form a VP is determined by invariant principles except the relative order of merged elements. These elements must be ordered, and if UG provides no specification, a parametric option is created. It seems that the content of the parametric option is simply to force a consistent choice on a grammar. I will return to this point in §3.4 and §3.5. It is an open question whether all postulated parameters of UG can be seen in this light, although this is a conceptually attractive

possibility as it makes variation an inherent part of the system rather than an unexplained accretion.

Fourth, parameters are usually thought of as binary, either/or options. This of course relates to the previous two points. The importance of this idea is that, like other aspects of UG, parameters are discrete entities: here as elsewhere, clines, continua, squishes and the like are ruled out. This idea does not prevent us from postulating parameters with more than two values; such parameters can always be reconstrued as networks of binary parameters. (Again, I will say much more about networks and hierarchies of parameters in §3.4 and §3.5.)

The parameters of UG tell us what is variant, and by implication what is invariant, in UG. They do three things that are of considerable general interest. First, they predict the dimensions of language typology. In our example, this implies that all languages can be divided into VO or OV. The VO languages include, in addition to English, the Romance languages, Greek, the Bantu languages, Thai, and many Papuan languages. The OV languages include Latin, the Indic languages, the Dravidian languages, Japanese, Korean, the Turkic languages, and many Amerindian languages (see Dryer (2005a: 338–41) and §6). Parameters can thus play a central role in the classification of languages. An important facet of this idea is that parameters may be able to define clusters of covarying properties, of the type stated by implicational and other types of universal put forward by typologists and others. We come back to **implicational universals** in §1.6. For example, VO vs. OV order seems to correlate with the relative order of auxiliaries and main verbs. We observe that auxiliary-verb (AuxV) and VO pattern together in English, while in German VAux and OV pattern together (again, we limit our attention to non-finite auxiliaries in German for the sake of simplicity):

- (8) a. John **can visit** Mary. (AuxV)
 b. Johann wird Maria **besuchen können**. (VAux)

Such clustering of variant properties is of central importance for language typology, since it establishes that syntactic variation is non-random. I will explore this and other word-order correlations more in §1.6.

Second, parameters should predict aspects of first-language acquisition. As noted above, first-language acquisition of syntax consists largely, perhaps exclusively, in fixing the values of parameters. In that case, we expect to be able to observe the effects of this parameter-fixing process in the development of syntax. Intensive research on this topic over recent years

has provided some intriguing conclusions on this point, which we will review in §3.1. (See Guasti (2002) for a much more detailed summary.) If variant properties cluster, as mentioned above, then it may be that one aspect of a grammar is acquired ‘for free’ once another is acquired (for example, AuxV order as a consequence of VO order). If this idea can be maintained, the task of the language acquirer is further simplified, and acquisition and typology are bound together.

Third, and of most concern to us here, parameters can tell us which aspects of syntax are subject to change in the diachronic dimension. On the basis of the English and German examples in (6) and (7), we can see that the relative order of the verb and its direct object, as a parameter, may be subject to change. In fact, we observed this in the Introduction when we compared Ælfric’s English with present-day English. There we saw that, at least in OE subordinate clauses, direct objects precede their verbs (‘... *that we the tree not touch*’; ‘*though that ye of the tree eat*’), while of course such orders are not possible in present-day English. If variant properties cluster together, then the clear prediction is that, all other things being equal, they will change together. (We will look at the case of VAux and OV order in the history of English in §1.6.2 and §2.5.) The nature of parametric change will be a central focus of this book.

In this section we have seen the motivation for the notion of parameter in Universal Grammar, and, albeit in a fairly rough form, an example of a parameter. The rest of this chapter is devoted to giving more detailed examples of parameters, both in the synchronic and the diachronic domains.

1.2. The null-subject parameter

The first parameter we will look at is the null-subject parameter. As with all the parameters to be discussed in this chapter, we first present the motivation for the parameter in the synchronic dimension, and then present the diachronic corollary.

1.2.1. *The null-subject parameter in the synchronic dimension*

The basic fact motivating the postulation of this parameter is that certain languages allow finite clauses not to express a definite, referential,

pronominal subject. In other languages, this is impossible. The contrast is illustrated by the following Italian and English examples:

- (9) a. Parla italiano.
b. *Speaks English.

We can observe that Spanish and Greek, among many other languages, pattern like Italian, while French patterns like English ((10c) is ungrammatical as a declarative, although it would be a well-formed imperative):

- (10) a. Habla español.
b. Mila ellinika.
c. *Parle français.

Thus Italian, Spanish, and Greek are null-subject languages, while English and French are non-null-subject languages. For now I will present parameters in the form of yes–no questions, and so the null-subject parameter may be informally stated as follows:

- A. Does every finite clause require an overt subject?
YES: non-null-subject languages (French, English ...).
NO: null-subject languages (Italian, Spanish, Greek, Japanese, Navajo ...)

We could rather literal-mindedly think of the language-acquirer asking itself² these questions about the linguistic data it is exposed to, as part of the process of fixing parameter values. I will return to the question of how to formulate parameters in §3.5. Equivalent to the formulation in A, we could present the parameter as an assertion ('Every finite clause has an overt subject'), and consider the value of the parameter to be the truth-value of this assertion. Both this formulation and that in A capture the binary nature of parameters automatically.

The null-subject parameter refers, as stated above, to finite, discourse-neutral clauses, and involves the interpretation of the null subject as a definite, referential pronoun. Many non-null-subject languages, including English, allow null subjects under other conditions. For example, both English and French extensively allow or require the subject of non-finite clauses to be null:

² I will refer to language acquirers using the neuter gender, since for our purposes the acquirer is really a 'relation between input data and a sequence of parameter values' (Clark and Roberts 1993: 303 and *passim*). Niyogi (2004: 12–13) similarly defines a language-learning algorithm as a map from a set *D* of data-tokens to a grammar *g*, and we may construe the grammar as a set of parameter values.

- (11) a. [(Him) smoking] bothers me.
b. John expects [(Mary) to leave soon].
c. Jean a essayé [de ... partir].
John has tried [... to leave].

Such subjects have somewhat different properties from the null subjects of (9) and (10), in that in (11b, c) the empty subject of the infinitive must be coreferent with the subject of the main clause (this is **subject control**) and in (11a) must be arbitrary. Accordingly, they have been analysed in a different way from those of (9) and (10). Most of the textbooks mentioned in the Introduction present the standard technical analyses of these constructions.

Both English and French also allow null subjects in special discourse environments or registers. Haegeman (2000: 130) gives the following examples from what she calls ‘written abbreviated registers’ (‘written registers in which pressures of economy seem to overrule the “core” grammar’ (132), including diaries, short notes and some kinds of colloquial speech):

- (12) a. ... cried yesterday morning.
(Plath 1983: 288)
b. Elle est alsacienne. ... Paraît intelligente.
She is Alsatian. ... Seems intelligent.
(Léautaud 1989: 48)

Such null subjects, in addition to being restricted to certain types of discourse and/or register, have certain special properties which distinguish them from the canonical null subjects of (9) and (10) (see Haegeman (2000: 138–41) for details). I will leave these cases aside here, and briefly return to them in §3.1.

Some languages allow expletive, or non-referential, null subjects, but not referential ones. German is one such language. In (13a) the expletive pronoun *es* is obligatorily ‘dropped’, while in (13b) the same pronoun in the same syntactic position, only now with a referential interpretation, cannot be dropped (examples from Cardinaletti (1990: 5–6)):

- (13) a. Gestern wurde (*es) getanzt.
yesterday was (it) danced
‘Yesterday there was dancing.’
b. Gestern war *(es) geschlossen.
yesterday was (it) closed
‘Yesterday it was closed.’

Owing to this restriction on their null subjects, such languages are not usually regarded as ‘full’ null-subject languages. Rizzi (1982: 143) in fact identified

what he called two ‘related but autonomous parameters’: one concerns whether, in our terminology, an empty pronoun is allowed at all, and the other whether it is allowed to be referential. In languages like English, both parameters are negative, while in Italian both are positive. In German, the first is positive and the second negative. Hence German allows non-referential null subjects, as in (13a), but not referential ones.

It is important to bear in mind that overt subject pronouns are allowed in finite clauses in null-subject languages, although they tend to have what we may loosely call an emphatic interpretation (this is indicated by putting the English pronoun in capitals in the translations below). Thus, alongside (9a) and (10a, b) we have:

- (14) a. Lui parla italiano. (Italian)
 HE speaks Italian
 b. Él habla español. (Spanish)
 HE speaks Spanish
 c. Aftos mila ellinika. (Greek)
 HE speaks Greek

This aspect of the interpretation of overt pronominal subjects in null-subject languages emerges slightly more clearly in examples such as (15). Here the overt pronoun in the adverbial clause does not allow the interpretation in which it corresponds to the subject of the main clause (see Vanelli, Renzi and Benincà 1985: 164):³

- (15) a. Il professore ha parlato dopo che (lui) è arrivato. (Italian)
 the professor has spoken after that (he) is arrived
 ‘The professor spoke after he arrived.’

³ In complement clauses, the same effect can be observed, although the interpretation is marginally possible rather than impossible (Luigi Rizzi p.c.):

- (i) Gianni dice che (lui) è il migliore.
 John says that (he) is the best

The same appears to be true in Greek (Anna Roussou, p.c.).

If the pronoun is stressed, modified or coordinated in the subordinate clause, coreference is possible (Cardinaletti 2003):

- (ii) a. Mario ha detto che LUI verrà domani.
 Mario has said that HE will-come tomorrow
 b. Mario ha detto che solo lui verrà domani.
 Mario has said that only he will-come tomorrow
 c. Mario ha detto che lui e sua madre verranno domani.
 Mario has said that he and his mother will-come tomorrow

- b. I Maria jelase apou (afti) idhe ton Yianni. (Greek)
 the Mary laughed after (she) saw Yiannis
 ‘Mary laughed after she saw Yiannis.’

In other words, the overt pronoun of the adjunct does not show the same ambiguity as its English and French counterparts in (16), in that it strongly prefers the interpretation which is disjoint from ‘the professor’, while the English and French pronouns are, out of context, ambiguous between this interpretation and the one where they correspond to ‘the professor’:

- (16) a. The professor spoke after he arrived.
 b. Le professeur a parlé après qu’il est arrivé. (= (16a))

These interpretative differences appear to be related to the null-subject parameter. Two further properties are held to go along with the positive value of the null-subject parameter, creating a cluster of properties of the type discussed in the previous section. The first of these is the possibility of ‘free inversion’, a construction in which the overt subject may follow the main verb, and bear a focused interpretation, i.e. as being new information. The construction is illustrated by the following Italian examples, while their French counterparts illustrate the impossibility of this construction in non-null-subject languages:⁴

- (17) a. È arrivato Gianni.
 b. *Est arrivé Jean.
 has arrived John
 ‘John has arrived.’
- (18) a. Hanno telefonato molti studenti.
 b. *Ont téléphoné beaucoup d’étudiants.
 have telephoned many students
 ‘Many students have telephoned.’

⁴ Here I illustrate ‘free inversion’ with two different types of intransitive verbs (**unaccusative** *arrivare* (‘arrive’) and **unergative** *telefonare* (‘telephone’); the distinction between the two types of intransitive is described in §2.3.1). Free inversion is heavily restricted with transitives in Italian. This is not the case in Spanish and Greek, where VSO order is readily available with transitives as shown by the following Greek example:

- (i) Episkevase o Yiannis ton ipolojisti mou.
 repaired the John the computer my
 ‘John repaired my computer.’
 (Roussou and Tsimpli 2006)

This shows that further parametric distinctions need to be made among null-subject languages.

The second property concerns the fact that the subject of a finite clause cannot be questioned if the **complementizer** introducing the clause is present. This constraint holds of English and French, as the following examples show:

- (19) a. *Who did you say that – wrote this book?
 b. *Qui as-tu dit qu’ – a écrit ce livre? (= (19a))

Here the questioned constituent (*who/qui*) corresponds to the subject of the subordinate clause, so there is a ‘gap’ in that position. We can think of the gap corresponding to the questioned constituent (*who/qui*) as a silent copy of that element, i.e. deriving from a structure in which *who/qui* is merged in the subject position, directly expressing the grammatical function of that element. According to this analysis, a subsequent operation, known as **Move**, places *who/qui* at the beginning of the sentence. We will look at Move in more detail in the next section (see Box 1.1) and at wh-movement in particular in §1.5.1. In earlier versions of the theory, copies were seen as ‘traces’ of movement, so the ungrammaticality of (19a) is sometimes known as the ‘complementizer-trace effect’. The notion that the presence of the complementizer determines the ungrammaticality of such examples is supported by the fact that (19a) becomes grammatical if *that* is omitted. In French, (19b) can be rendered grammatical by altering the form of the complementizer from *que* to *qui*. These points are illustrated in (20):

- (20) a. Who did you say – wrote this book?
 b. Qui as-tu dit qui – a écrit ce livre? (= (20a))

In null-subject languages, on the other hand, it appears that complementizer-trace effects are not found. The subject of a finite clause introduced by a complementizer can readily be questioned:

- (21) a. Chi hai detto che – ha scritto questo libro? (Italian)
 who have-2sg said that – has written this book
 b. Pjos ipes oti – egrapse afto to vivlio? (Greek)
 who said-2sg that – wrote this the book
 ‘Who did you say wrote this book?’

Rizzi (1982) conjectured that free inversion and the absence of complementizer-trace effects in null-subject languages are really two cases of the same phenomenon. His proposal was that the subject *chi/pjon* in (21) does not correspond to a copy following the complementizer, as indicated there, but rather to a copy *in the inverted position*. This makes it possible to maintain

that there may be a universal (not parameterized) ban on copies occupying the position immediately following a complementizer. We would naturally want to reduce this to a more fundamental and general aspect of UG, but that is a separate issue. Striking confirmation for Rizzi's conjecture comes from certain Northern Italian dialects. These dialects at first sight seem not to be null-subject languages, in that a subject **clitic** (SCL), if available, is obligatory. I illustrate with examples from the Florentine dialect (see Brandi and Cordin (1989)):

- (22) a. *Parla.
 Speaks
 b. *Mario parla.
 Mario speaks
 c. Mario e parla.
 Mario SCL speaks
 'Mario speaks.'
 d. E parla.
 SCL speaks

Brandi and Cordin, following Rizzi (1986b), show that the subject clitics are really agreement markers, comparable to the verbal endings, rather than pronouns. In that case, these varieties in fact are null-subject languages, since in examples like (22d) they show subject agreement (in the form of the clitic) but no actual overt subject. Now, where the subject is freely inverted, the subject clitic has a special neutral form and the verb is always 3sg, rather than agreeing with the inverted subject (similarly, the past participle does not agree in cases where it would do so in Standard Italian; compare (23b) and (24b)):

- (23) Florentine:
- a. **Gli** ha telefonato delle ragazze.
 SCL has telephoned some girls
 'Some girls phoned.'
- b. **Gli** è venuto delle ragazze.
 SCL is come-masc.sg. some girls
 'Some girls came.'
- (24) Standard Italian:
- a. Hanno telefonato delle ragazze.
 have-3pl phoned some girls
 'Some girls phoned.'
- b. Sono venute delle ragazze.
 are-3pl come-fem.pl. three girls
 'Some girls came.'

Florentine does not allow the Standard Italian agreement patterns, as (25) shows:

- (25) a. *Le hanno telefonato delle ragazze.
 SCL-3pl.f. have-3pl phoned some girls
 ‘Some girls phoned.’
 b. *Le son venute delle ragazze.
 SCL-3pl.f. are-3pl. come-f.pl. some girls
 ‘Some girls came.’

When the subject is questioned, only the agreement pattern in (23) is allowed:

- (26) a. Quante ragazze gli ha parlato con te?
 how-many girls SCL has talked with you
 ‘How many girls have talked to you?’
 b. Quante ragazze gli è venuto con te?
 how-many girls SCL is come with you
 ‘How many girls have come with you?’

This extends to examples where the subject of a finite clause is questioned:⁵

- (27) a. Quante ragazze tu credi che gli abbia parlato?
 how-many girls you think that SCL have-3sg talked
 ‘How many girls do you think talked?’
 b. Quante ragazze tu credi che è sia venuto?
 how-many girls you think that SCL be-3sg come
 ‘How many girls do you think have come?’

The above examples are all from Brandi and Cordin (1989: 112–27). As they point out, ‘when the subject is questioned, agreement patterns as if inversion has taken place’ (124). This confirms Rizzi’s earlier proposal for Standard Italian.

The parameter in A thus relates together the following three properties:

- (28) a. The possibility of a silent, referential, definite subject of finite clauses;
 b. ‘Free subject inversion’;
 c. The absence of complementizer-trace effects.

These properties can be related together in one of two ways. On the one hand, we could assume that the subject position is universally present, and hence that it is filled with a silent pronoun – designated *pro* – in null-subject languages. The null-subject parameter then basically states that some

⁵ In (27a), *gli* could be interpreted as a masculine dative clitic ‘to him’, but that is not relevant here. In (27b) *è* is a subject clitic. It is not clear what underlies the allomorphic variation here.

languages have *pro*, while others do not. In free inversion contexts, and where an inverted subject is questioned in the presence of a complementizer, a non-referential *pro* occupies the obligatory subject position. This is the analysis put forward in Rizzi (1986a).

On the other hand, we could assume that in null-subject languages the verb itself, or more precisely the verbal agreement, is able to contain the subject argument. We need appeal neither to a universally required subject position, nor to a silent pronoun filling it in null-subject languages. The null-subject parameter relates directly to properties of verbal agreement, and it has often been noticed that null-subject languages tend to have ‘rich’ subject-agreement inflection, in many cases marking each person-number combination with a distinct ending. This is true in Standard Italian and Modern Greek, for example. If the subject function is marked by verb agreement, a separate expression of the subject is not necessary. If a separately expressed subject does appear, it is free to appear in a range of positions in the clause depending on its discourse function (topic, focus, etc.). This view was originally put forward by Borer (1986) and has been recently restated in the context of minimalist technical assumptions by Barbosa (1995) and Alexiadou and Anagnostopoulou (1998).

I will not choose between the two analyses of the null-subject parameter here. The purpose of this section is to introduce the phenomenon, and to illustrate the clustering of the properties in (28). A number of other properties have been proposed as part of this cluster, notably the possibility of ‘clitic-climbing’ whereby a clitic pronoun dependent on an infinitive may appear in a superordinate clause if that clause contains a verb of the relevant class. This is possible in Italian and Spanish, but not in French (although Cinque (2004) takes a different view of the situation in French):

- (29) a. Gianni lo vuole mangiare.
 John it wants to-eat
 ‘John wants to eat it.’
 b. *Jean le veut manger.
 John it wants to-eat

See Kayne (1989) on this.

Another property that may be correlated with the null-subject parameter is the ability of an infinitive to appear in a position preceding a clitic pronoun. Again, this is possible in Italian but not in French (except in imperatives):

- (30) a. Parlargli sarebbe un errore.
 b. *Parler-lui serait une erreur.

to-speak to-him would-be a mistake
 ‘It would be a mistake to speak to him.’

See Kayne (1991: 648ff.) for discussion. How these last two features may be related to the analyses of the null-subject parameter proposed by Rizzi and Borer is unclear, and I will not speculate about that here.

Having seen some aspects of the null-subject parameter in the synchronic domain, i.e. how it can define certain types of cross-linguistic variation, it is now time to look at a case where its value appears to have changed over time.

1.2.2. *The null-subject parameter in the diachronic dimension: changes in the history of French*

In the previous section, we gave French as an example, along with English, of a non-null-subject language. In fact, it appears that the value of the null-subject parameter has changed at least once in the history of French. Here we will see that, at earlier stages of its history, French was a null-subject language. The value of this parameter seems to have changed in approximately 1600.

In Old French (OF, 842–1300) and Middle French (MidF, 1300–1500) we can readily find examples of null subjects, such as the following:

(31) Old French:

- a. Tresqu'en la mer **cunquist** la tere altaigne.
 until the sea conquered-3sg the land high (*Roland*, 3)
 ‘He conquered the high land all the way to the sea.’
- b. Si **chāi** en grant povreté.
 thus fell-1sg into great poverty (*Perceval*, 441)
 ‘Thus I fell into great poverty.’
- c. Si en **orent** moult grant merveille.
 thus of-it had-3pl very great marvel (*Merlin*, 1)
 ‘So they wondered very greatly at it.’
 (Roberts 1993a: 124ff.)

(32) Middle French:

- a. Et ly **direz** que je me recommande humblement a elle...
 and her will-say-2pl that I myself recommend humbly to her
 ‘And you will say to her that I humbly ask her good will ...’
 (S 131, 16)
- b. Ne vous **pourroye** a demi dire le tresgrant dueil.
 neg you could-1sg at half say the very-great grief
 ‘I could not tell you half the great grieving.’ (S 165, 17)
 (Vance 1997: 260)

The relevant subject pronouns are conspicuously absent in these examples. As we saw in the previous section, such pronouns must be present in Modern French finite clauses. In other words, if we perform the exercise carried out with Ælfric's English and Modern English in the Introduction, and update all aspects of the examples in (32) except for their syntax, the result is ungrammatical Modern French, viz.:

- (32') a. *Et lui **direz** que je me recommande humblement à elle ...
 b. *Ne vous **pourrais** à moitié dire le très grand deuil.

We can see that the grammar of French has changed historically; structures which were formerly grammatical no longer are. This is a clear indication that a parameter has changed.

Earlier stages of French also show free inversion:⁶

- (33) a. Tant fu de bone hore nez li chevaliers! (Q 10, 12)
 so-much was of good hour born the knight
 'The knight was born at such a propitious hour!'
 b. car assez l'ot eschaufé li serpenz (Q 95, 1)
 for much him-had warmed-up the snake
 'for the snake had heated him up considerably'
 (Vance 1997: 77)

Again, these examples yield ungrammatical sentences if 'translated' word-for-word into Modern French:

- (33') a. *autant fut au bon moment né le chevalier!
 b. *car beaucoup l'eut rechauffé le serpent

Once more we see clear evidence of parametric change in the history of French.

French does not appear to have allowed complementizer-trace effects. Sentences of this kind are rather rare, and the main studies of OF and MidF word order do not comment on them. Nevertheless I am unaware of any clear examples analogous to (21) in OF or MidF. If, as we shall see

⁶ Vance (1997: 86ff.) argues that OF examples of this type are not really cases of free inversion of the Italian type. Whilst it is true that there are clear differences in comparison with Italian, which Vance amply illustrates, we have already seen that there are differences among null-subject languages regarding the properties of the free-inversion construction – see note 4. Hence I retain the idea that these are cases of free inversion, while recognizing that the term 'free inversion' probably does not designate a unitary phenomenon. The fact that the canonical subject position is not occupied by an overt subject in examples like (33) indicates that this construction is connected to the null-subject parameter.

below, null subjects were limited to main clauses, this is what we in fact would expect, as complementizer-trace phenomena are by definition phenomena found only in embedded clauses.

On the other hand, the other properties which may pattern with the possibility of null subjects – clitic-climbing and infinitive-clitic order – are found in earlier stages of French. Clitic-climbing appears to die out during the seventeenth century as a fully productive option (Ayres-Bennett 2004: 209–19), while infinitive-clitic orders are found at early periods of OF:

- (34) a. Nous **lui** devons rendre gloire.
 we to-him must give glory
 ‘We must give him glory.’
 (Calvin, cited in Roberts (1994: 233))
- b. et y mist serganz a plenté pour **garder** le
 and there put sergeants a-plenty to guard him
 ‘and put many sergeants there to guard him’
 (*Helcanus* 253; de Kok 1985: 115)

So it seems clear that French used to be a null-subject language. Can we determine when the value of the parameter changed? This appears to have happened around 1600 (see Fontaine (1985), Roberts (1993a: 204ff.), Vance (1997: 321ff.), Sprouse and Vance (1999)). The remarks of a contemporary grammarian, Maupas, in his 1607 *Grammaire française*, are interesting in this respect. Maupas points out that subject pronouns are omitted in three main contexts that are of interest. These are (i) where the subject is 1pl or 2pl, (ii) after certain conjunctions, notably *et* (‘and’) and *si* (‘thus, so’), and (iii) where the subject is a non-referential. Examples of null subjects of these types are as follows:

- (35) a. J’ay receu les lettres que – m’avez envoyees.
 I have received the letters that me-have-2pl sent
 ‘I have received the letters you have sent me.’
- b. Il vous respecte et si – vous servira bien.
 he you respects and so you will-serve well
 ‘He respects you and will serve you well.’
- c. Rarement – advient que ces pronoms nominatifs soient omis.
 Rarely happens that these pronouns nominative be omitted
 ‘It rarely happens that these nominative pronouns are omitted.’
 (Roberts 1993a: 215–6)

(Example (35c) is from Maupas’ own discussion of null subjects.) Although null subjects are still permitted at this point, the range of possible contexts is highly restricted, and became still more restricted in the seventeenth

century. In seventeenth-century French, only non-referential null subjects like that in (35c) are found, and these become progressively limited to fixed expressions. It is reasonable to conclude, then, that French ceased to be a fully null-subject language around 1600.⁷

What we have said so far implies that OF and MidF were like Modern Italian or Spanish in the relevant respects. However, there is a significant complication in French, in that null subjects are sensitive to the type of clause in which they appear. Null subjects are much more widely attested in main clauses than in embedded clauses (see Price (1971); Einhorn (1974); Foulet (1990); Vanelli, Renzi, and Benincà (1985); Adams (1987a, b); Vance (1988; 1997); Roberts (1993a)). Examples such as the following are of great interest in this connection:

- (36) a. Ainsi s'acorderent que il prendront par nuit.
 thus they-agreed that they will-take by night
 'This they agreed that they would take by night.'
 (*Le Roman du Graal*, B. Cerquiglini (ed.), Union Générale d'Éditions, Paris, 1981: 26; Adams (1987b: 1); Roberts 1993a: 84)
- b. dont la joye fut tant grant par la ville qu'elle
 of-which the joy was so great in the town that-it
 ne se pourroit compter
 neg self could count
 'the joy concerning which was so great around town that it could not be counted'
 (*Jehan de Saintré* 160, 5; Sprouse and Vance 1999: 263)

⁷ It is possible that null expletives have survived in Modern French in one construction, Stylistic Inversion (see Kayne and Pollock (1978; 2001), Pollock (1986)). This construction looks similar to Italian free inversion (see Pollock (1986) on this) but occurs only in certain contexts: questions, relatives, exclamatives, clefts (i.e. 'wh' contexts) and subjunctives. Here is an interrogative example:

- (i) A qui a téléphoné Marie?
 to whom has phoned Marie
 'Who did Marie call?'

It can be observed that the preverbal subject position is empty. This construction may thus be a survival of expletive null subjects in one set of highly restricted contexts (see Roberts (1993a: 217–19); Vance (1997: 172ff.)). Kayne and Pollock (2001) suggest that examples like (i) contain a null subject which 'doubles' the postverbal subject, a silent version of *elle*. Their analysis implies that Modern French retains referential null subjects in certain contexts. The contrasts between (32) and (32'), (33) and (33') nevertheless clearly show that a major change in the distribution of null subjects has taken place in the history of French. Here we begin to see that the null-subject parameter may be rather more complex than first appears. Below we will see further indications of this.

In (36a) we observe a null 3pl subject in the main clause, and an overt 3pl subject in the complement clause (*il* could be plural in OF; the verb form is clearly 3pl here). The context clearly favours an interpretation where the overt pronoun in the subordinate clause corresponds to the null subject in the main clause. We observed in the previous section that this is generally not possible in null-subject languages (see (15) and the comment in note 3 above). Similarly, in the MidF example (36b) the overt pronoun *elle* in the adverbial clause corresponds to *la joye* in the main clause, a further case of the same type. If we suppose, however, that null subjects were only allowed in main clauses in OF and MidF, then the subordinate clauses in (36) would be non-null-subject contexts. The overt pronouns would be required, much as in Modern French or English, whether or not they correspond to a main-clause element. The apparent anomaly in (36) as compared to what we observed in connection with Italian and Greek in (15) can then be seen as connected to the fact that OF and MidF null subjects are largely restricted to main clauses. In the next section, we will see that this restriction is related to the fact that the finite verb occupies a different position in main and embedded clauses in these periods of French (a further difference with the modern language).⁸

In the discussion of Florentine in the previous section, I mentioned that, despite the obligatory appearance of subject clitics, Florentine is nevertheless a null-subject language since these clitics are not pronouns but agreement markers. Thus a sentence containing just a subject clitic and a verb such as *e parla* (SCL speaks, i.e. 'he speaks') in fact features a null subject. If this is true for Florentine examples of this type, could it be true for Modern French examples such as *il parle*? Here, the pronoun *il* is certainly a subject clitic in that it is a phonological dependent of the verb: no material except for other clitics can intervene between it and the verb, and it cannot be stressed, co-ordinated, or modified. These characteristics were originally noticed in Kayne (1975) and are subject to close scrutiny and an interesting theoretical interpretation in Cardinaletti and Starke (1999).

Certain authors (among them Jaeggli (1982); Roberge (1990); Sportiche (1998)) have argued that there is little or no difference between French and Northern Italian dialects regarding the status of subject clitics. Without going

⁸ In fact, in MidF the differences between main and embedded clauses regarding both null subjects and verb position are less marked than in OF. See Vance (1997) for detailed discussion and analysis of this.

into the details of this debate, it is possible that at least some varieties of Modern French treat the subject pronouns as agreement markers. One particularly interesting case is Algerian French, as reported in Roberge and Vinet (1989). In this variety, ‘doubling’ of the subject is ‘extremely frequent’ (Roberge and Vinet 1989: 53), as indeed in many varieties of modern colloquial French. Hence we find examples of the following type (here the absence of a comma after the first word indicates that they are to be read without an intonational break at this point, i.e. with a single intonation contour):

- (37) a. Marie elle vient.
 Mary she comes
 b. Pierre il mange.
 Peter he eats

Nevertheless, unlike in Florentine and many other Northern Italian dialects, the doubling is not obligatory. There are also Northern Italian dialects where doubling is not obligatory. Some, mainly Veneto varieties, seem to optionally allow a subject clitic to co-occur with an overt subject argument, for example:

- (38) Nane (el) magna.
 John SCL eats
 ‘John eats.’

But (38) is ambiguous between ‘John, he eats’ and ‘John eats’. That is, we cannot tell whether ‘John’ is the topic and ‘he’ is the subject or ‘John’ is the subject and ‘he’ is the subject clitic marking agreement (topics do not have to be subjects, cf. *Beans, John likes*). We can tell the two options apart if we replace *John* with a negatively quantified argument, as such arguments cannot be topics (cf. the oddity of an English example like **Nobody, he eats beans* as opposed to *John, he eats beans*). Doing this, we find that the subject clitic is not allowed in Veneto:

- (39) Nisun (*el) magna.
 nobody (*SCL) eats
 ‘Nobody eats.’

This suggests that the subject clitic in (38) may be a subject; its apparent optionality is due to the option of interpreting *Nane* as either the topic or the subject. This in turn would imply that, in Veneto, subject clitics are in complementary distribution with argumental subjects. In other words, Veneto seems just like French in this respect. There are good reasons, nevertheless, to consider the Veneto clitics as agreement markers rather

than pronouns. In particular these elements follow the preverbal negative marker, an option which is impossible for subjects. See Poletto (2000) for detailed discussion of this and other tests for the structural status of Northern Italian subject clitics, and Cardinaletti and Repetti (2003) for an opposing view, which classifies Veneto with Modern French. But the question remains open regarding Standard French.

Many Northern Italian dialects allow the equivalent of (39), however. According to Poletto (2000: 142), this is the situation in 'Friulian, most Piedmontese dialects and some Ligurian and Lombard varieties'. The sentence in (40) is a Friulian example:

- (40) Nisun al mi capiss.
 no-one SCL me understands
 'Nobody understands me.'
 (Friulian: S. Michele a T. Poletto 2000: 142)

In such varieties there is no doubting the status of the subject clitic. It can only be an agreement marker, as it co-occurs with what is quite clearly a subject argument. Now, Algerian French allows examples like (40):

- (41) Personne il sait que c'est leur mère.
 no-one he knows that it's their mother
 (Roberge and Vinet 1989: 53)

In this variety of French, then, we are led, by parity of reasoning with what was stated regarding the Northern Italian dialects, to conclude that the subject clitics are agreement markers rather than pronouns, although their presence is optional when an overt subject is present. The consequence of this is that an example like *il parle* features a null subject. Hence Algerian French is a null-subject language, whatever the status of Standard French.

It may be, then, that French has developed, or is developing, as follows:

- (42) Stage I: null subjects (in main clauses), strong subject pronouns;
 Stage II: no null subjects, weak subject pronouns;
 Stage III: null subjects, subject clitics.

In connection with Stage I, we can observe that in OF and MidF, subject pronouns were not phonologically dependent on the verb in that they could be separated from the verb, modified and co-ordinated:

- (43) a. Et il, a toz ses oz, s'en ala.
 and he, and all his army, went away
 (Price 1971: 145; Roberts 1993a: 113)

- b. Se je meïsmes ne li di ...
if I self not him say ...
'If I don't tell him myself ...'
(Franzén 1939: 20; Roberts 1993a: 114)
- c. Et qui i sera? Jou et tu.
and who there will-be? I and you
'And who will be there? I and you.'
(Price 1971: 145; Roberts 1993a: 113)

These are the characteristics of 'strong' pronouns (see Cardinaletti and Starke (1999)).

By the sixteenth century, these pronouns had 'weakened' and, as we have seen, null subjects disappeared at just about this time. This gave rise to the Stage II system. Stage III is what we find in Algerian French, whether or not other varieties have reached this stage. Here the subject clitics have further 'weakened' to become agreement markers and the value of the null-subject parameter changes again. Although it is rather approximate, and raises a variety of questions, the diachronic development in (42) illustrates the interactions between subject pronouns and the null-subject parameter. Since the Medieval Northern Italian dialects had null subjects in main clauses, like OF (see in particular Vanelli, Renzi, and Benincà (1985) on this), we may conjecture that these dialects have also gone through the stages in (42); see also Poletto (1995) on null subjects and subject pronouns in Renaissance and seventeenth-century Veneto, and Renzi (1983) on eighteenth-century Florentine.

The above discussion shows that the null-subject parameter has changed at least once in the recorded history of French and of many Northern Italian dialects. This is then our first example of parametric change.

1.3. Verb-movement parameters

Starting with the seminal work of den Besten (1983), Emonds (1978; 1980), and Pollock (1989), a great deal of attention has been paid to parameters involving the overt position of the verb in the clause. Since, as we stated in the Introduction, we take verbs universally to merge with their complement(s) to form a VP, the natural way to analyse variation in the surface position of the verb is by postulating a further operation of movement which places the verb in some other position in the clause after it has formed the VP. We will see that there are two basic varieties of verb-movement,

which can be distinguished in terms of the target of movement, i.e. the position to which the verb is moved. As in the previous section, I first present the parametric variation from a synchronic point of view, and then move on to the evidence for parametric change.

1.3.1. *Verb-movement in the synchronic dimension*

1.3.1.1. Verb-movement to T

If, first, V merges with its direct object to form a VP as described in the Introduction, and, second, the grammatical function ‘direct object’ is defined in terms of Merge in this way (as we briefly mentioned in §1.1), and, third, if Merge is strictly binary, as is usually assumed, then it follows that verbs are always adjacent to their direct objects. To a good first approximation, this is true for English. The main verb and its direct object are generally adjacent, with adverbial and other material such as negative markers unable to intervene between them.⁹

- (44) a. *John **kisses often** Mary.
 John **often kisses** Mary.
 b. *John **eats not** chocolate.
 John does **not eat** chocolate.

In many other languages, however, this straightforward situation does not obtain. In French, finite verbs are naturally separated from their direct objects by adverbs of various kinds, while the negative element *pas* obligatorily intervenes between the finite verb and its object:

⁹ One construction where this is not true is the so-called ‘double-object’ construction, where the indirect object must intervene between the verb and the direct object:

- (i) a. John sent Mary a bunch of flowers.
 b. *John sent a bunch of flowers Mary.

Here it is the notional indirect object that must be adjacent to the verb:

- (ii) a. John often sends/*sends often Mary flowers.
 b. John does not send/*sends not Mary flowers.

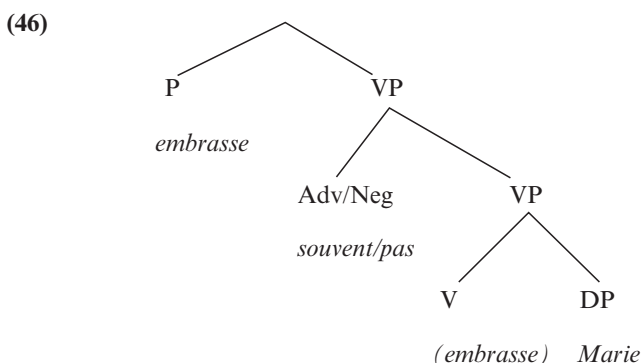
The ‘dative alternation’ relates (ia) to (iii):

- (iii) John sent a bunch of flowers to Mary.

This alternation has given rise to much discussion in the literature on syntactic theory. Most of the textbooks cited in the Introduction devote some space to this question, and so I will leave it aside here. We will come back to this construction again briefly in §2.3.

- (45) a. Jean **embrasse souvent** Marie. (=44a)
 *Jean **souvent embrasse** Marie.
 b. Jean (ne) **mange pas** de chocolat. (=44b)
 *Jean (ne) **pas mange** de chocolat.

Beginning with Emonds (1978), two ideas have been put forward which together describe what is going on in French examples of this type. First, it is assumed that adverbs like *often* and *souvent* as well as the negative elements *pas* and *not* are merged in positions to the left of VP (positions whose precise nature need not detain us; Pollock (1989: 378–9) shows that negation and adverbs in fact occupy different positions – see note 10). Second, it is proposed that in French the finite verb is required to move to a position still further to the left, outside VP. Very roughly, then, we have a structure like the following for (45a), where *embrasse* moves to some position P:



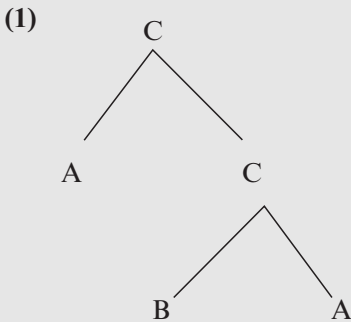
What is the position P merged with VP here? We can get a clue, perhaps, by comparing English examples containing auxiliaries with (44) and (45):

- (47) a. John **has often** kissed Mary.
 b. John **has not** kissed Mary.

The natural position for *often*, and the only position for *not*, is between the auxiliary and the main verb. Now, auxiliaries mark, among other things, tense, aspect, and mood. Observe also that in the grammatical version of (44b), the auxiliary *does*, whose sole content is that of tense/agreement marker, must precede *not*; in fact, auxiliary *do* is in complementary distribution with the auxiliary *have* of (47). If auxiliaries carry tense information, then we might think that there's a special position for tense-markers outside VP to the left of the position of adverbs and negation.

BOX 1.1 Technical aspects of movement

A brief word on some technical aspects of movement is perhaps required here. Chomsky (2001) suggests that movement is really second Merge, or internal Merge. That is, when movement happens, Merge applies to an element that has already been merged once, and merges it again, as shown in (1):



Here A has first merged with B, and is subsequently remerged with C. Since Merge builds structure, second Merge places the moved element ‘higher’ in the tree than the position of first Merge. In fact, movement generally ‘extends the tree’ in this sense. (Structure is always built up leftwards, i.e. the right branches are generally the recursive branches, for reasons put forward in Kayne (1994), which we will briefly review in §2.5.4).

When an element undergoes movement, a copy remains in the original position. Usually, this copy is not pronounced – it undergoes deletion in the phonological component. I have indicated the copy of the moved verb here in brackets in the main text, and will continue to follow this practice. In the examples involving questioning of the subject of a finite clause of the type discussed in the previous section we thus have the following:

- (2) a. Who did you say (who) wrote this book?
 b. Chi hai detto che *pro* ha scritto questo libro (chi)?

Assuming Rizzi’s (1986a) analysis of null subjects, there are two different types of silent category in (2b).

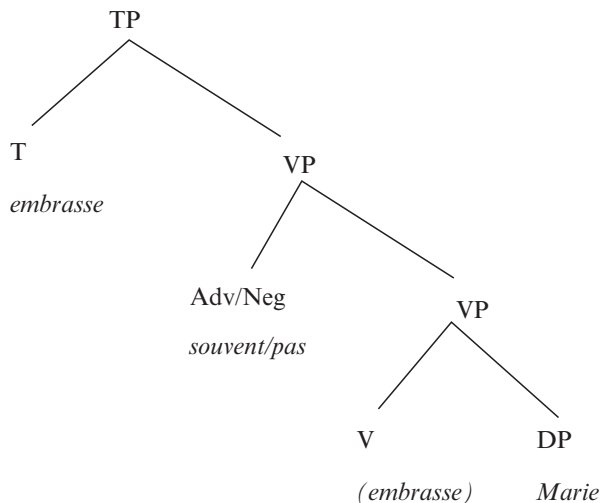
What causes movement to take place? Since its incidence varies cross-linguistically, as this section is intended to show, it must be controlled by a parametrically varying property. We can for present purposes regard the property of triggering movement (or functioning as an attractor for some category) as a property arbitrarily associated with different positions in different grammatical systems. In §2.5.4, I will introduce a notation which encodes this arbitrary, cross-linguistically variant property of certain positions.

Moreover, we have mentioned that the verb-movement in French examples like (45) only affects finite verbs. Infinitives, for example, cannot precede *pas*:¹⁰

- (48) a. **Souvent embrasser** Marie, ...
 often to-kiss Mary ...
 b. Ne **pas embrasser** Marie, ...
 Neg not to-kiss Mary ...
 c. *N'**embrasser pas** Marie, ...

There appears to be a connection between the target of finite-verb movement and the temporal specification of the clause. Accordingly, let us identify P with the category Tense, i.e. the category of morphemes (including auxiliaries) or features whose content determines the temporal specification of the clause. For simplicity, we can think of this as the position T, associated with features such as [Present], [Past], etc., although the reality is certainly considerably more complex. When T merges with VP, the resulting category is a TP, a temporal expression. So we can replace (46) with (49):

(49)

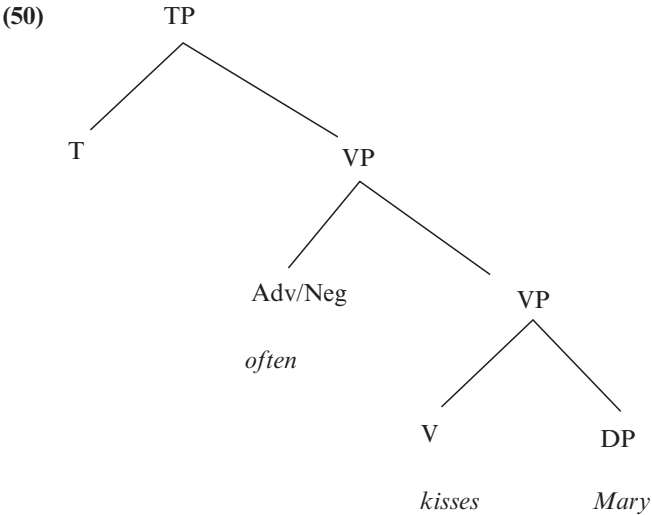


¹⁰ Pollock (1989: 379) points out that infinitives can in fact precede *souvent*, but not *pas*. The logical conclusion is that there are two targets for movement, one above *pas* and one between *pas* and *souvent*:

(i) ... X *pas* Y *souvent* V ...

Here X corresponds to P in (43). Finite verbs obligatorily move to X = P in French, as we have said, while infinitives move optionally to Y. See Pollock for details, and Belletti (1990); Cinque (1999) for refinements.

The structure for English examples like the grammatical version of (44b) is as follows:¹¹



Comparing (49) and (50), we can observe that English and French differ in the position occupied by the finite main verb. More precisely, finite main verbs move to T in French, while their English counterparts do not. So we formulate the following parameter:

- B. Does V move to T in finite clauses?
 YES: French, Welsh, Italian, Icelandic, Greek ...
 NO: English, Swedish, Danish, ...

The difference in the value of the parameter in B accounts for the word-order differences between English and French that we have observed here, and similar differences among the other languages listed in B. The basic difference is that languages with a positive value for this parameter allow the order *V-Adv/Neg-object* while those with a negative value for it require the order *Adv/Neg-V-object*.

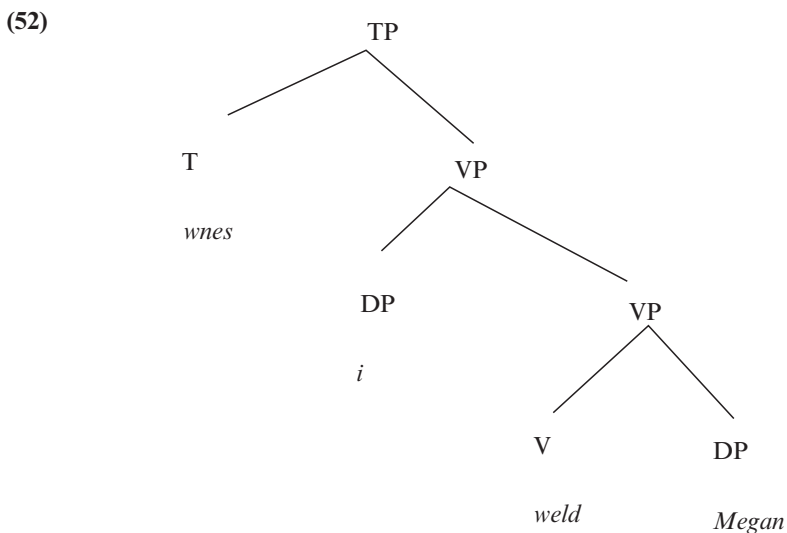
A particularly interesting case in this connection is Welsh, which we can in fact take to be representative of the Celtic VSO languages generally, and perhaps of other VSO languages such as Classical Arabic and Biblical

¹¹ Here T contains the feature [Present], I assume, which lacks a phonological realization. The relation between T and the inflection on the unmoved verb is presumably mediated by the operation Agree, which I will introduce in the next section.

Hebrew.¹² The usual order in finite clauses in Welsh is VSO, as (51) illustrates:

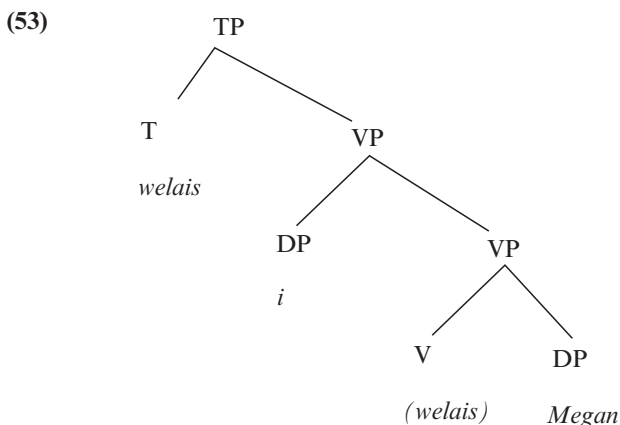
- (51) a. Fe/mi **welais** Megan.
 PRT saw I Megan
 ‘I saw Megan.’
 b. Fe/mi **wnes** i weld Megan.
 PRT did I see Megan
 ‘I saw Megan.’

By definition, in VSO order the verb is not adjacent to the direct object. If we are to retain the assumptions about the structural relation of the verb and its direct object described above, we must assume that the verb moves ‘over’ the subject in simple VSO sentences like (51a). This analysis is supported, in the case of Welsh, by the existence of an alternative way of expressing almost any sentence whose verb is in a simple tense by using a construction involving the auxiliary *gwneud* (‘do’) and a non-finite form of the verb (the so-called ‘verbal noun’), as in (51b). Here the auxiliary precedes the subject, which in turn precedes the verb, which in turn precedes the direct object. So we have the order AuxSVO. If auxiliaries are in T, as we suggested for English, then we can assign a structure like (52) to (51b); the ‘particle’ *fe/mi* is merged with TP in a higher position, and so we leave it out of (52):



¹² The type of analysis to be described almost certainly does *not* generalize to all languages which have surface VSO order, however. See the papers in Carnie and Guilfoyle (2000), especially Lee (2000); Massam (2000); Rackowski and Travis (2000). For an in-depth study of Hebrew and Palestinian Arabic dialects, see Shlonsky (1997).

Here the subject is merged with the ‘core’ VP containing the verb and the direct object. Examples without auxiliaries and with VSO order involve V-movement of the type we observed above for French, as follows:



VSO order thus involves V-movement to T. What distinguishes Welsh from French is a further parameter concerning whether the subject appears in VP or in a higher position, which we may take to be the analogous position in relation to TP, viz.:

(54) [_{TP} Jean [_{TP} embrasse souvent Marie]].

Here the internal structure of the inner TP is that diagrammed in (49). So Welsh subjects appear in VP, as in (52, 53), and French subjects appear in TP as in (49) – here we presumably have a further parameter.

If subjects are universally defined in terms of their merged position, like direct objects, then we can take them to be universally merged with VP as in the putative Welsh structures in (52, 53) and moved in languages like French (and English) to the position shown in (54). (We will modify this idea in §2.3.1.) In that case, VSO order of the Welsh type arises where the subject does not move and the verb does.¹³ Positions of the type occupied

¹³ Things are slightly more complex in fact, since there is evidence that in Welsh and in some other Celtic languages (perhaps all of them) the subject does move from the VP-internal position, although it does not move as far as it does in English and French. The evidence comes from the fact that the subject must precede the negative element (*d*)*dim*:

- (i) (Ni) **ddarllenodd** Emrys **ddim** o'r llyfr.
 (Neg) read Emrys not of the book
 'Emrys didn't read the book.'

by the subjects in both French and Welsh, which we can generalize as in (55a), are known as specifiers. Positions formed by simple merger with a head, such as that occupied by direct objects, are known as complements. Finally, the category formed by the head and the complement is conventionally labelled X' (so the inner TP in (54) should really be T')

- (55) a. [_{XP} YP [_{X'} X ...]] (YP is a specifier of XP)
 b. [_{X'} X YP] (YP is the complement of X)

Since Merge is binary, there can be only one complement per head. On the other hand, there is no restriction in principle on the number of specifiers. We can now define 'direct object' as the complement of V and 'subject' as the specifier of V. The parameter distinguishing French from Welsh concerns whether the subject moves to a specifier of TP ('SpecTP') – this is what we have in French – or whether it stays in a specifier of VP (or at least in a specifier position structurally lower than SpecTP – see note 13). XP-movement always creates specifiers, since it must always 'extend the tree' (see Box 1.1). Head-movement, on the other hand, always targets other head-positions.

Here we have seen two cases of V-to-T-movement. In SVO systems, V-to-T-movement can distinguish systems like English from those like French (incidentally, this shows that SVO order does not result from a unique parameter-setting). V-to-T movement can play a central role in deriving VSO order in VSO languages, if we allow for the possibility that the subject might be realized relatively 'low' in the structure in such languages. (It is very likely that VSO does not represent a single parameter-setting either, as mentioned in note 12.)

1.3.1.2. V-movement to C: full and residual V2

A further parameter of verb-movement distinguishes verb-second (V2) languages from languages like English, French, and Welsh. In V2 languages, as the name implies, the finite verb is the second element in the clause. More precisely, it follows exactly one XP. German is perhaps the best known example of a V2 language, and we can illustrate the V2 phenomenon with the following German examples:

The best way to ensure this is to place negation outside VP (unlike the structure given in (50), but in line with what was implicit in note 10) and then cause the subject to move to a position in between T and negation.

- (56) a. Ich **las** schon letztes Jahr diesen Roman.
I read already last year this novel
- b. Diesen Roman **las** ich schon letztes Jahr.
this novel read I already last year
- c. Schon letztes Jahr **las** ich diesen Roman.
already last year read I this book
'I read this novel last year already.'
- d. *Schon letztes Jahr ich **las** diesen Roman.

In each of the grammatical examples, one XP precedes the finite verb *las*: the subject *ich* in (56a) (which is a DP, although it consists of just one word); the direct object *diesen Roman* in (56b), and the adverbial phrase *schon letztes Jahr* in (56c). It is impossible to 'stack up' more than one XP before the finite verb, as the ungrammaticality of (56d) shows. This should be compared with English and French examples like the following:

- (57) a. Last year I read this novel.
b. L'an dernier j'ai lu ce roman. (= (57a))

It is easy to see that in (56a) and (56c) at least the verb is separated from the direct object and has thus moved out of VP. In (56b) the direct object has moved – let us refer to the operation which places an XP in first position in V2 clauses as topicalization. So in this example both the verb and the direct object have moved out of VP.

It is clear that V2 only applies to finite verbs; infinitives and participles are unaffected by it. The counterparts of (56) in a periphrastic tense show just the auxiliary in second position:

- (58) a. Ich **habe** schon letztes Jahr diesen Roman **gelesen**.
I have already last year this novel read
- b. Diesen Roman **habe** ich schon letztes Jahr **gelesen**.
this novel have I already last year read
- c. Schon letztes Jahr **habe** ich diesen Roman **gelesen**.
already last year have I this book read
'I read this novel last year already.'

These examples also show that non-finite verbs follow their direct objects in German, as we briefly saw in §1.1, and will see once more in §1.5. Since it affects finite main verbs and auxiliaries, V2 seems to involve the T-position. The question is: does it involve movement just to the T-position or does it involve movement from the T-position to a still higher position?

It is usually thought that in V2 clauses the verb moves to a higher position than T (but see note 15). The reason for this is that V2 is largely restricted to main clauses in many languages, with German again being a typical case. In embedded clauses, the finite verb or auxiliary must appear in final position in German:

- (59) Du weißt wohl,
You know well
a. ... daß ich schon letztes Jahr diesen Roman **las**.
... that I already last year this novel read
b. ... daß ich schon letztes Jahr diesen Roman **gelesen habe**.
... that I already last year this book read have
- (60) Ich frage mich,
I ask myself
a. ... ob ich schon letztes Jahr diesen Roman **las**.
... if I already last year this book read
b. ... ob ich schon letztes Jahr diesen Roman **gelesen habe**.
... if I already last year this book read have

Embedded clauses are typically introduced by a complementizer: *daß* ('that') in (59) and *ob* ('if/whether') in (60). Let us label the category complementizer C. We can then see that C merges with TP to form an embedded clause:

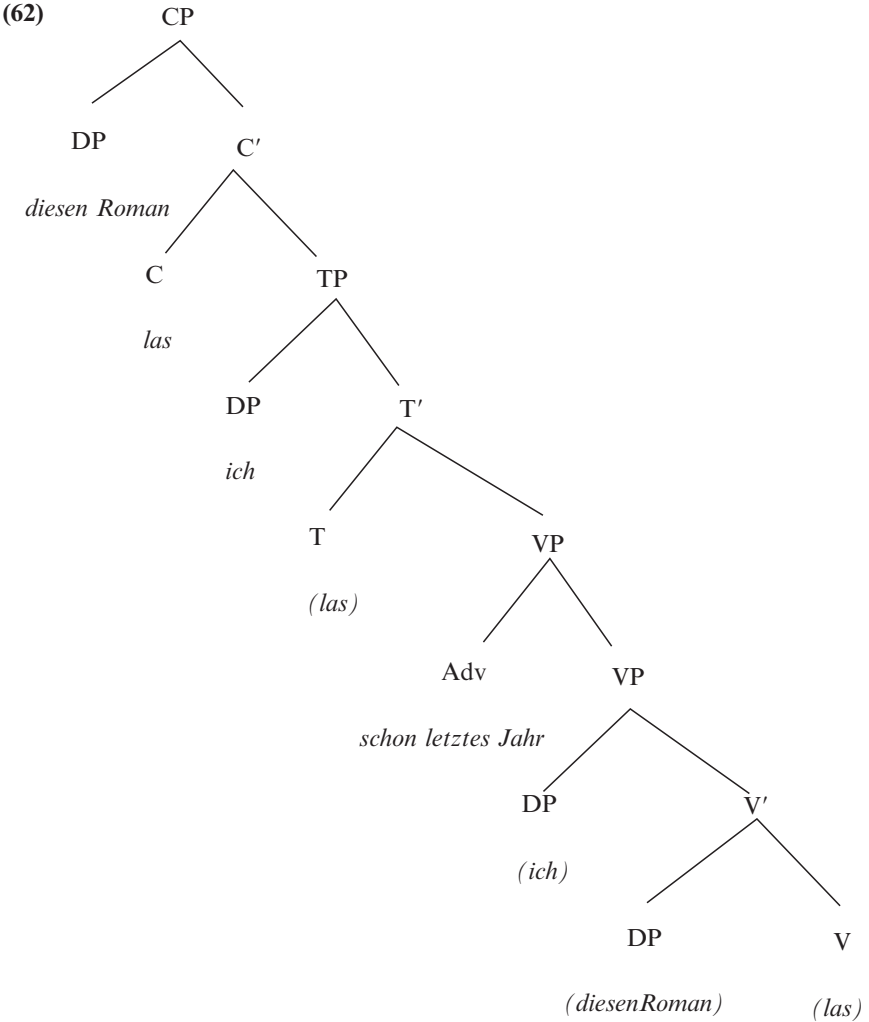
- (61) [_{CP} ob/daß [_{TP} ich schon letztes Jahr diesen Roman las]]

English embedded clauses like *whether/that I read this novel last year* would also have the structure in (61) at the CP-level. Now, we can understand the ban on V2 in embedded clauses if we take it that, first, the verb moves to C in V2 clauses and, second, since C already contains a complementizer in embedded clauses, the position is already filled and so the verb is unable to move there. This idea was first proposed by den Besten (1983). If the verb moves to C in V2 clauses the structure of an example like (58b) will be as shown in example (62):¹⁴

Here we see that the topicalized XP occupies a Specifier of CP (see (55a)). We can thus postulate the following parameter:

- C. Does the finite verb move to C in finite main clauses?
YES: German, Dutch, Swedish, Icelandic, Danish, Kashmiri, Romansch ...
NO: English, French, Italian, Welsh ...

¹⁴ In the lowest VP, the copy of the direct object precedes V, an instance of the general OV pattern of German. I return to this in §1.6 and §2.5.



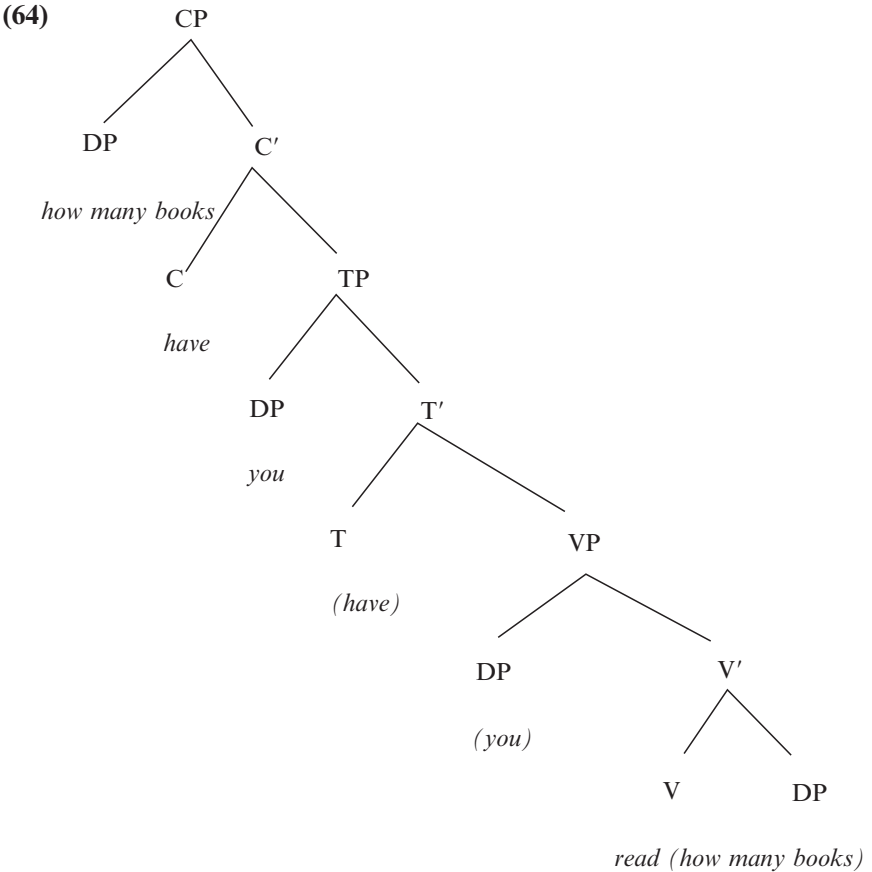
Of course, this parameter is connected with only one aspect of V2, the verb-movement part. Why and how movement of just one XP to SpecCP is required is a separate matter that I will not go into here.¹⁵

¹⁵ It should also be pointed out that many analysts do not accept that V moves to C in all types of V2 clause. Travis (1984) and Zwart (1997) argue that V does not move beyond T in subject-initial V2 clauses such as (56a) and (58a). See Schwartz and Vikner (1996) for a defence of the view presented in the text. In some languages, notably Yiddish and Icelandic, the restriction to main clauses is much less rigid, which has led to the suggestion that V does not move to C in these – so-called ‘symmetric V2’ – languages (see Vikner (1995: 80–7) for discussion). Müller (2004a) offers a rather different analysis of V2, which does not involve verb-movement combined with XP-movement but rather movement to SpecCP of a single larger category which contains only XP and V.

The two parameters just presented interact in various ways. Both English and French have been referred to as ‘residual V2’ languages, since a finite verb or auxiliary may occupy C in just some kinds of main clauses. In particular, we find this construction in main-clause interrogatives in both languages:

- (63) a. How many books have you read?
 b. Combien de livres as-tu lus? (= (63a))

Here the thing to notice is that *have/as* precedes the subject. We are assuming that auxiliaries are T-elements, and we saw earlier that subjects occupy SpecTP in English and French, and so here the auxiliaries must move to a higher position. The natural candidate is C, given the analysis of V2 in (62). This gives rise to the structure in (64) for (63a):



The idea that the auxiliaries move to C in (64) is supported by the fact that indirect questions, where a complementizer is present, do not allow this order:

- (65) a. *I wonder if has he read that book.
 b. *Je me demande si a-t-il lu ce livre. (= (65a))

Now, observe that main verbs can move to C in main-clause questions in French, but this possibility is limited to auxiliaries in English:

- (66) a. Quel genre de livre préfères-tu lire en vacances? (= (66b))
 b. *What kind of book prefer you to read on holiday?

Instead of (66b), where no other semantically motivated auxiliary is present, the dummy auxiliary *do* must be used:

- (67) What kind of book do you prefer to read on holiday?

We can understand the fact that only auxiliaries can move to C in residual V2 clauses in English in terms of parameter B introduced earlier, combined with the following restriction on head-movement:¹⁶

- (68) The Head-Movement Constraint:
 In a single step of movement, a head can only move as far as the next head-position up.

Example (68) means that the verb cannot move directly from V to C in a structure like (62), ‘skipping’ an intermediate head-position like T. Now, in English, main verbs do not move to T; we saw above that English has the negative value for parameter B. Thus English main verbs cannot move to C: they cannot move there directly because of (68), and they cannot move there via T because of the negative value of parameter B. On the other hand, auxiliaries, which I continue to assume are merged as T, may move to C in English, and French main verbs may move to C via T thanks to French having the positive value of parameter B.

¹⁶ (68) is an instance of an important general condition on movement, which requires that any instance of movement target the nearest ‘available’ position. The precise structural definition of ‘nearest’ will emerge in the next section. The definition of ‘available’ is rather complex and will not detain us here. Again, these matters are dealt with in detail in the textbooks cited earlier. See also Rizzi (1990; 2000).

By adopting (68), we can understand the difference between English and French observed in (66) as a consequence of the different values of parameter B in the two languages. Taken together, then, the verb-movement parameters account for considerable range of cross-linguistic variation, which we can summarize in Table 1.1.¹⁷

All of the major Germanic, Romance, and Celtic languages, as well as Greek, appear to fall into the typology created by the three parameters we have discussed in this section. (I have included two null-subject Romance languages, alongside Greek, also a null-subject language.) We can also deduce the properties of certain languages which we have not yet discussed. For example, Breton combines (a form of) V2 with VSO (see Shafer (1994);

Table 1.1 Synchronic verb-movement parameters

	Parameter B	Parameter C	Subject moves to SpecTP
German, Dutch, Icelandic	Yes	Yes	Yes
Swedish, Danish, Norwegian	No ¹⁸	Yes	Yes
French, Italian, Spanish, Greek	Yes	No	Yes
Breton	Yes	Yes	No
English	No	No	Yes
Welsh, Irish	Yes	No	No
(unattested)	No	Yes	No
(unattested)	No	No	No

¹⁷ The last two possibilities in Table 1.1 may be unattested because T must always be a target for movement, of a head, an XP or both. This rather mysterious property of T is captured in some versions of the Extended Projection Principle of Chomsky (1982; 1995). See §2.5.4 for more on the (rather different) conception of the Extended Projection Principle (or EPP) in recent minimalist theory.

¹⁸ (71a) below shows that Danish, like Swedish and Norwegian, lacks V-to-T movement in non-V2 clauses, since the verb is adjacent to the direct object in VP and follows *ikke* ('not'). These are nevertheless V2 languages. The same applies in Dutch and German if VP follows T, as shown in (62). However, if VP precedes T, then it may be that V moves to T in these languages – see note 30. To comply with (68), V must pass through T 'on its way' to C in V2 clauses in these languages. That English does not allow this possibility in examples like (66b) may be connected either with the existence of the dummy auxiliary *do* in English (an item with no counterpart in the Scandinavian languages) or with the precise formulation of the way in which C triggers V-movement in full V2 languages. I will leave this matter open here.

Jouitteau (2005)). Here we see how parameters may combine to define a range of cross-linguistic options.

1.3.1.3. Further properties related to verb-movement

Two further properties may be related to verb-movement. First, Holmberg (1986) observed a phenomenon in the Scandinavian languages which has come to be known as ‘object shift’. This can be observed in Danish examples like the following (from Vikner (1995)):

- (69) a. Hvorfor **læste** studenterne **den** ikke?
 why read students-the it not
 ‘Why didn’t the students read it?’
 b. Hvorfor **læste** studenterne ikke **artikeln**?
 why read students-the not articles-the
 ‘Why didn’t the students read the articles?’

Example (69) shows that pronouns appear before the negation in V2 clauses, while non-pronominal objects follow the negation. Taking negation in Danish to occupy a position similar to English *not* and French *pas* (see (49)), we see that the object has been moved out of VP in (69a). The non-pronominal object in (69b), on the other hand, presumably remains in VP (these objects may move in Icelandic, but not in Swedish, Norwegian, or Danish). Object shift does not take place where the verb does not move; for example in subordinate clauses – (70a) – or where V is non-finite, as in (70b):

- (70) a. Det var godt at han ikke **købte den**.
 it was good that he not bought it
 ‘It was good that he didn’t buy it.’
 b. Hvorfor skal studenterne ikke **læse den**?
 why shall students-the not read it
 ‘Why don’t the students have to read it?’

We thus arrive at what has become known as Holmberg’s generalization: the object moves only if the verb moves.¹⁹ Object shift is therefore connected to a positive value for one of the two parameters we have seen in this section.

Second, Bobaljik and Jonas (1996: 228–9) propose that ‘transitive-expletive constructions,’ illustrated in (71), are found just in languages which have a positive value for parameter B:

¹⁹ Holmberg (1999) has reformulated this generalization, but I retain his earlier formulation here for expository reasons.

- (71) a. Es essen einige Mäuse Käse in der Küche. (German)
 there eat some mice cheese in the kitchen
 ‘There are some mice eating cheese in the kitchen.’
- b. Er hat iemand een appel gegeten. (Dutch)
 there has someone an apple eaten
 ‘Someone has eaten an apple.’
- c. Það hafa margir jólasveinar borðað þúðing.
 there have many Christmas-trolls eaten pudding
 ‘Many Christmas trolls have eaten pudding.’
 (Icelandic: Bobaljik and Jonas (1996: 209))

Compare the situation in English and the Mainland Scandinavian languages, which have the negative value for parameter B:

- (72) a. *There has someone eaten an apple.
 b. *Der har nogen spist et æble. (Danish)
 there has someone eaten an apple
 c. *Det har någon ätit ett äpple. (Swedish)
 there has someone eaten an apple

Bobaljik and Jonas restrict their attention almost exclusively to Germanic languages, and it is unclear exactly how transitive expletive constructions might manifest themselves in a null-subject language or a VSO language; I therefore leave this question aside.

Having seen two important parameters regarding verb-movement, and a number of assumptions regarding the structure of the clause and the nature of movement, I now turn to the diachronic evidence that the values of these parameters can change.

1.3.2. *Verb-movement in the diachronic dimension*

Here we will see evidence that the two verb-movement parameters introduced in the previous section have changed at different periods in the history of English. Parameter C has also changed in the history of French, and this latter change may be connected to the change in the value of the null-subject parameter which we discussed in §1.2.2 above.

1.3.2.1. V-to-T in earlier English

In earlier English, until approximately 1600, main verbs were able to move to T. Warner (1997: 381–6) provides a very interesting discussion of the

chronology of this change; see also the discussion of Kroch (1989) in §4.1.5. We can see this from the fact that main verbs could be separated from their direct objects by negation and by adverbs, as in the following examples:

- (73) a. if I **gave not** this accompt to you
 ‘if I didn’t give this account to you’
 (1557: J. Cheke, Letter to Hoby; Görlach 1991: 223; Roberts 1999: 290)
- b. The Turkes . . . **made anone redy** a grete ordonnaunce.
 ‘the Turkes . . . soon prepared a great ordnance.’
 (c1482: Kaye, *The Delectable Newsse of the Glorious Victorye of the Rodyans agaynest the Turkes*; Gray 1985: 23; Roberts 1993a: 253)

Examples like these have a slightly familiar ‘Shakespearean’ feel for many speakers of present-day English. Shakespeare lived from 1564 to 1616, and so in his English V-movement to T was possible; hence examples of the type in (74) can be found in his plays and poems. Despite this air of familiarity, the examples in (73) are ungrammatical in present-day English. This shows us that parameter B has changed since Shakespeare’s time.

We take examples like (73) to tell us that sixteenth-century English had the ‘French’ value for parameter B. If so, then following the reasoning outlined at the end of the previous section, we expect that main verbs were able to move to C in residual V2 environments at this time. This is correct, as (75) shows:

- (74) What **menythe this pryste?**
 what does this priest mean
 (1466–7: Anon., from J. Gairdner (ed.), 1876, *The Historical Collections of a London Citizen*; Gray 1985: 11; Roberts 1993a: 247)

Example (75) is ungrammatical for modern speakers. Here the main verb moves from V to T to C. V-to-T movement is allowed owing to the positive setting of Parameter B, as in Modern French. Compare (74) with the French example in (66a); we take the two sentences to be structurally isomorphic in relevant respects.

Also, given Holmberg’s generalization as presented in the last section, we expect to find object shift in sixteenth-century English. Again, this expectation is borne out:

- (75) a. if you **knew them** not
 (1580, John Lyly; Roberts 1995: 274)
- b. They **tell vs** not the worde of God.
 (1565, Thomas Stapleton; Roberts *ibid.*)

Here V has moved to T; we know this because it precedes *not*. The pronominal object (in fact it is an indirect object in (75b) – see note 8) also precedes *not* and so we take this element, too, to have left VP. In (75b), the direct object presumably remains within VP; see again note 9.

Transitive expletive constructions are also found in earlier periods of English, up until approximately the sixteenth century, as (76) shows:

- (76) a. Within my soul there doth conduce a fight.
 (Shakespeare; Jonas (1996: 151))
 b. ... there had fifteene severall Armados assailed her.
 (1614; Raleigh Selections 151; Jonas 1996: 154)

So we witness the clustering of non-adjacency of the main verb and its direct object, main-verb-movement to C in residual V2, transitive expletive constructions and object-shift in sixteenth-century English. This cluster of properties in sixteenth-century English, and the absence of these properties in present-day English, can be described by postulating a change in the value of parameter B at some point between the sixteenth century and the present.

1.3.2.2. V2 in diachrony

At a still earlier period, until approximately the fifteenth century (see Fischer *et al.* (2000: 132–7) for discussion), English had the positive value of parameter C. In other words, OE and ME were V2 languages. The following examples illustrate V2 in OE:

- (77) a. Se Hæland **wearð** þa gelomlice ætiwed his leornung-cnihtum.
 the Lord was then frequently shown his disciples
 ‘The Lord then frequently appeared to his disciples.’
 (*ÆCHom* I, 15.220.21; Fischer *et al.* 2000: 106)
 b. On twam þingum **hæfde** God þæs mannes sawle gegodod.
 in two things had God this man’s soul endowed
 ‘With two things had God endowed the man’s soul.’
 (*ÆCHom* I, 1.20.1; Fischer *et al.* 2000: 107)
 c. þa **astah** se Hælend up on ane dune.
 then rose the Lord up on a mountain
 ‘Then the Lord went up on a mountain.’
 (*ÆCHom* I, 12. 182.1; Fischer *et al.* 2000: 108)

In (77a), the subject precedes the finite auxiliary; in (77b) a PP precedes the auxiliary; in (77c) the adverb *þa* (‘then’) precedes the finite verb.

Just like the other modern Germanic languages (with the exceptions given in note 15), V2 is only found in main clauses in OE. The verb is typically final in embedded clauses (although this word order is less consistent in OE than in Modern Dutch or German – see Fischer *et al.* (2000: 145ff.); §1.6.2; and §2.5 for more detailed discussion of this point):

- (78) a. ... þæt ic þas boc of Ledenum gereorde to Engliscre
 ... that I this book from Latin language to English
 spræce **awende**.
 tongue translate
 ‘... that I translate this book from the Latin language to the
 English tongue’
 (*AHTH, I*, pref, 6; van Kemenade 1987: 16)
- b. ... þæt he his stefne up **ahof**.
 ... that he his voice up raised
 ‘... that he raised up his voice.’
 (*Bede* 154.28; Pintzuk 1991: 77)
- c. ... forþon of Breotone nædran on scippe lædde **wæron**.
 ... because from Britain adders on ships brought were
 ‘... because vipers were brought on ships from Britain.’
 (*Bede* 30.1–2; Pintzuk 1991: 117)

The general picture is thus strikingly similar to Modern German, as a comparison with the previous section and the brief discussion of German word order in §1.1 reveals.

In ME, at least until the latter part of this period, V2 is still consistently observed:

- (79) a. On þis gær **wolde** þe king Stephne tæcan Rodbert.
 in this year wanted the king Stephenb seize Robert
 ‘During this year King Stephen wanted to seize Robert.’
 (c12; *ChronE* (Plummer) 1140.1; Fischer *et al.* 2000: 130)
- b. Oþir labur **sal** þai do.
 other labour shall they do
 (*Ben. Rule(I)* (Lnsd) 33.20; Fischer *et al.* 2000: 131)
- c. Nu **loke** euerich man toward himseluen.
 now look every man to himself
 ‘Now it’s for every man to look to himself.’
 (c13; *Ken.Serm* 218.134; Fischer *et al.* 2000: 130)

In (79a) an adverbial PP precedes the finite auxiliary; in (79b), the direct object precedes the finite auxiliary, and in (79c) an adverb precedes the finite verb. It appears then that both parameters B and C have changed in the recorded history of English.

French, too, was a V2 language at an earlier stage of its history, as the following OF examples illustrate:

- (80) a. Par Petit Pont **sont** en Paris entré.
by Petit Pont are (they) in Paris come

BOX 1.2 Verb second in Old English

There is a major complication as regards V2 in OE, first pointed out by van Kemenade (1987, Chapter 4). It appears that pronouns can intervene between the first constituent and the verb in what would otherwise be a V2 clause. This is illustrated in (1a) with a subject pronoun and in (1b) with an object pronoun:

- (1) a. Hiora untrymnesse **he sceal** rowian on his heortan.
their weakness he shall atone in his heart
'He shall atone in his heart for their weakness.'
(*CP* 60.17; Pintzuk 1999: 136)
- b. Pin agen geleafa **þe hæfþ** gehæledne.
thy own faith hee has healed
'Thine own faith has healed thee.'
(*BlHom* 15.24–5)

Furthermore, where the initial constituent is a *wh*-phrase, the negative element (*ne*) or the adverb *þa* (very roughly 'then'), the verb was required to follow this constituent directly, and thus precede a pronoun. This is illustrated in (2):

- (2) a. Hwæt **sægest þu,** yrþling?
what sayest thou, earthling
'What do you say, ploughman?'
(*ÆColl.*, 22.23; Fischer *et al.* 2000: 40)
- b. Ða **weard he** to deofle awend.
then was he to devil changed
'Then he was changed to a devil.'
(*AHTh*, I, 12; van Kemenade 1987: 138)
- c. Ne **worhte he** þeah nane wundre openlice.
nor wrought he yet no miracles openly
'But he worked no miracles openly.'
(*AHTh*, I, 26; van Kemenade 1987: 138)

The correct analysis of this construction has been and remains the subject of much debate: see van Kemenade (1987, Chapter 4); Pintzuk (1991); Roberts (1996); Fischer *et al.* (2000: 106–8); Fuß and Trips (2002); Haerberli (2002) for discussion and different types of analysis. I will leave this matter aside here, although understanding the nature of the contrast between (1) and (2) is clearly of great importance for understanding correctly the syntax of verbs and pronouns in OE.

The same effect is observed in southern dialects of ME, but not, according to Kroch and Taylor (1997), in northern dialects.

‘They entered Paris by the Petit Pont.’

(*Charroi de Nîmes*, 27; Roberts 1993a: 95)

- b. Sa venoison **fi**st a l’ostel porter.
 his goods made (he) to the hostel carry
 ‘He had his goods carried to the hostel.’

(*Charroi de Nîmes*, 27; Roberts 1993a: 95)

- c. Li cuens Guillelmes **fu** molt gentix et ber.
 the count G. was very kind and good
 ‘Count G. was very kind and good.’

(*Charroi de Nîmes*, 28; Roberts 1993a: 95)

In (80a) the PP *par Petit Pont* precedes the verb, while in (80b) the direct object *sa venoison* occupies the preverbal position. In (80c), the subject precedes the finite verb. Embedded clauses generally did not allow V2 order in OF, although this may not have been true in early OF, prior to the thirteenth century – see Adams (1988a, b, c); Dupuis (1988, 1989); Hirschbuhler and Junker (1988); Hirschbuhler (1990); Roberts (1993a: 132ff.); Vance (1997: 162–6). Vance (1997: 133) shows that subject-initial order is found in 46 per cent of main clauses in her thirteenth-century prose text (*La Queste del Saint Graal*) and in 88–96 per cent of embedded clauses, depending on the precise nature of the clause in question. This period of OF, at least, appears therefore to feature movement of the finite verb to C quite generally in main clauses. We conclude that OF had the positive value for parameter C.

In MidFr, the nature of V2 changes somewhat, but V2 main clauses are still found, as the following examples illustrate. In (81a), the conjunction *et* can be thought of as being outside the V2 clause; it does not ‘count’ as the initial constituent in the V2 clause:

- (81) a. Et aussi **fi**s je de par vous.
 and also did I from by you
 ‘And I did likewise with respect to you.’
 (S 104, 15; Vance 1997: 266)

- b. Si **suis** je aussi bien armé.
 thus am I as well armed
 ‘Thus I am as well armed.’
 (S 290, 22; Vance 1997: 267)

We saw in the previous section that OF was a null-subject language. We also saw in the discussion of (36) that there is reason to think that null subjects were restricted to main clauses, unlike in Modern Italian, Spanish, and Greek. I suggested that this restriction on null subjects was related to verb positions. We are now in a position to state the observation more clearly: in OF, null subjects are allowed only in V2 clauses. This observation was first made by Thurneysen (1892), and has been restated and analysed in the context of generative approaches by Vanelli, Renzi, and Benincà (1985); Adams (1987a, b); Vance (1988; 1997); and Roberts (1993a). In terms of the analysis of V2 put forward in the previous subsection, we can say that the subject may be null just where the verb moves to C.

One advantage of linking V2 and null subjects in this way has to do with a class of embedded clauses which allows V2. These are known as the so-called complements to ‘bridge verbs’, verbs of saying, thinking, etc.²⁰ V2 is allowed in such clauses in all of the V2 Germanic languages, including those which do not otherwise allow V2 in embedded clauses (see (59), (60), and note 15). In German, the complementizer disappears and the verb must be in the subjunctive, at least according to prescriptive grammar, while in the other languages the V2 order follows the complementizer, illustrated here by Danish (from Vikner (1995: 71)):

- (82) a. Watson behauptete, dieses Geld hätte Moriarty gestohlen.
 Watson claimed this money had(subjunc) M. stolen
 b. Watson påstod at disse penge havde Moriarty stjålet.
 Watson claimed that this money had Moriarty stolen
 ‘Watson claimed Moriarty had stolen this money.’

²⁰ They are known as bridge verbs because a constituent in their complement clause can readily undergo wh-movement. Compare (ia), which features a bridge verb, with (ib), which does not:

- (i) a. Who does John think that Mary likes (who)?
 b. ?Who does John regret that Mary likes (who)?

The metaphorical notion is that verbs such as *think* in some way form a bridge across which the wh-constituent may move from its merged position in the embedded clause to the SpecCP position of the main clause.

However such clauses are to be analysed (and see Vikner (1995: 70ff.) for a survey of the possibilities; see also Penner and Bader (1995) on some semantic properties of this construction in certain varieties of German), we can make an immediate inference regarding OF if it is true that V2 and null subjects are connected in that language. The inference is this: if OF also allows V2 in complements to bridge verbs, then it should allow null subjects in exactly these embedded clauses. In (83a) I illustrate that embedded V2 is indeed possible in this context, and in (83b) I show that null subjects are also possible here (OF allowed the complementizer *que* ('that') to optionally drop, as in (83b)):

- (83) a. Et il respondirent que de ceste nouvele **sont** il moult lié.
 and they replied that of this news are they very happy
 'And they replied that they were very happy with this news.'
 (*Le Mort le Roi Artu*, J. Frappier (ed.), Droz, Genève, 1964: 45; Adams 1987b)
- b. Or voi ge bien, plains **es** de mautalant.
 now see I well, full are-2sg of bad-intentions
 'Now I see clearly that you are full of bad intentions.'
 (*Charroi de Nîmes* 295; Roberts 1993a: 97)

We expect to find (83b) if we have (83a) and if null subjects and V2 are linked in OF. So let us conclude that this is the case. If so, a diachronic prediction immediately emerges: we might expect the loss of V2 and the loss of null subjects to be connected in the history of French. The two changes are at least chronologically correlated, in that both took place in the sixteenth century (see Adams (1987a, b); Roberts (1993a); Vance (1997) for extensive discussion and analysis of this). This is then an interesting case where two putatively independent parameter values appear to be linked. A similar connection between V2 and null subjects is attested in the Medieval Northern Italian dialects discussed in the previous section in connection with the null-subject parameter; these varieties had the positive setting for both parameters, and both appear to have changed, with the subsequent reversal of the value of the null-subject parameter caused by the change in status of subject pronouns (as discussed in the last section).

To conclude this discussion, I present the diachronic correlate of the table given at the end of the previous subsection (the second line is the same as in the earlier table, as are lines 5 and 6; the latter two are however relevant diachronically in that earlier stages of these languages appear elsewhere in the table):

Where different stages of the same language appear in different lines of the table, this indicates that a parameter change has taken place. Parameter B changed in English ca. 1600; parameter C changed in English ca. 1400, in French

Table 1.2 Parameters of Verb-movement in older Germanic, Romance, and Celtic languages

	Parameter B	Parameter C	Subject moves to Spec TP
OE, ME, OF, MidFr	Yes	Yes	Yes
(Danish, Swedish, Norwegian)	No	Yes	Yes
ENE, Modern French	Yes	No	Yes
Middle Welsh ²¹	Yes	Yes	No
NE	No	No	Yes
Modern Welsh	Yes	No	No
(unattested)	No	Yes	No
(unattested)	No	No	No

ca. 1600 (and was apparently linked to a change in the null-subject parameter). Finally, parameter C changed in seventeenth-century Welsh (see Willis (1998)).

1.4. Negative concord

In this section, I want to present a rather different sort of parameter from the ones discussed up to now. This kind of parameter has to do with the presence or absence of a particular class of lexical items and how these items interact with syntactic principles and mechanisms, rather than being directly concerned with word order. It is also more connected with semantics, in that questions of how certain words and syntactic relations among words are semantically interpreted arise. Nevertheless, we will give it a formal characterization. There are probably a number of parameters of this sort, but the one we will concentrate on has to do with the expression of negation, and I will call it the negative concord parameter.

1.4.1. *Negative concord synchronically*

The basic observation concerning negative concord can be simply stated: in some languages, each morphologically negative expression corresponds to a **logical negation**, while in others, several morphologically negative expressions in a given syntactic domain may combine to express a single logical negation. Languages of the latter type are known as negative-concord languages; languages of the former type are non-negative-concord

²¹ Willis (1998) argues that Middle Welsh was a V2, VSO language, with the V2 parameter changing in the seventeenth century.

languages. English is a non-negative concord language, as the following set of examples, taken together, show:

- (84) a. I saw nothing.
 b. I didn't see anything.
 c. I didn't see nothing.

Examples (79a, b) each contain a single logical negation, and can be translated into a quasi-logical statement such as 'There is no x such that I saw x ', and a single negative morpheme: *no* or *nothing* in (84a) and *n't* in (84b). Combining the two expressions of negation in a single sentence, as in (84c), gives rise to a slightly awkward sentence, which clearly has a positive interpretation.²² In (Standard) English, then, two realizations of negation in a single sentence (or clause) give rise to a positive statement. Each negative expression 'counts' as an autonomous logical negation, and so the positive interpretation is a consequence of the logical truth $p \equiv \neg \neg p$.

In French, (84a, b) are both naturally translated as (85):

- (85) Je *n'ai* *rien* vu.
 I neg-have nothing seen
 'I have seen nothing/I haven't seen anything.'

Here, we observe that there are two negative morphemes, *ne* and *rien*. We can see that *rien* is intrinsically negative because in isolation it corresponds to 'nothing', unlike *anything*, which is ungrammatical in isolation:

- (86) a. Qu'est-ce que tu as vu? Rien!
 what is it that you have seen? Nothing
 'What have you seen? Nothing!'
 b. What have you seen? *Anything!

The same is true of the other negative words (or n-words) in French, such as those given in (87) below. The two negative expressions in (85) correspond to a single logical negation. This is an example of negative concord;

²² Many readers will recognize (84c) as the non-standard equivalent of (84b). This shows that non-standard varieties of English have negative concord. Therefore, if negative concord is determined by a parameter, we conclude that non-standard English and Standard English have at least one different parameter value and therefore are different grammatical systems. Labov (1972: 130–201) provides a detailed discussion of negative concord in one variety of non-standard English, which we will come back to in §4.2.1. The normative notion that it is illogical to interpret sentences like (84c) as containing a single negation reflects ignorance of the nature of negative concord on the part of normative grammarians, rather than ignorance of logic on the part of speakers of non-standard English.

the two expressions must agree, i.e. show concord, in the formal expression of a single logical negation.

French makes extensive use of negative concord, as the following examples show. I have systematically given the English translations using both *no* and *not ... any* (or the equivalent of *any*):

- (87) a. Jean *n'a* *jamais* dit cela.
 John neg-has never said that
 'John has never said that/John hasn't ever said that.'
- b. Jean *n'a* *aucun* espoir de gagner.
 John neg-has no hope to win
 'John has no hope of winning/John hasn't any hope of winning.'
- c. Jean *n'a* vu *personne*.
 John neg-has seen no-one
 'John has seen no-one/John hasn't seen anyone.'

What is the nature of the negative-concord relation? We can analyse it in terms of the relation of Agree, introduced in Chomsky (2000; 2001). Agree is a matching relation holding between formal features in a particular syntactic domain. Formal features are categorial features like V and N, as well as features like Person, Number, Gender, and Case, which are clearly relevant for morphosyntactic well-formedness. Some formal features, such as Number, have semantic content; others, such as certain Case features (for example, Nominative, Accusative) lack semantic content. Formal features which have a semantic interpretation are known as interpretable features; those which lack such an interpretation are uninterpretable. Chomsky argues that certain formal features may be interpretable in one position and uninterpretable in another: for example, Person and Number features are uninterpretable on verbs but interpretable on nouns. The Agree relation eliminates uninterpretable features, which is a necessary condition for a sentence to be grammatical. Interpretable features are not eliminated, as they are interpreted by the semantic component.

Agree is defined as follows (the Greek letters here stand for any syntactic category):

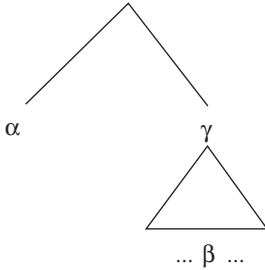
- (88) *a* Agrees with *β* where:
 (i) *a* and *β* have non-distinct formal features;
 (ii) *a* asymmetrically c-commands *β*.

We define asymmetric c-command in (89):

- (89) *a* asymmetrically c-commands *β* if and only if *β* is contained in the structural sister of *a*.

Two categories are structural sisters if they are ‘at the same level’ in a tree diagram. To put it another way, two categories α and β are structural sisters as a result of being merged with one another. Asymmetric c-command as defined in (89) thus obtains between α and β in the following schematic configuration:

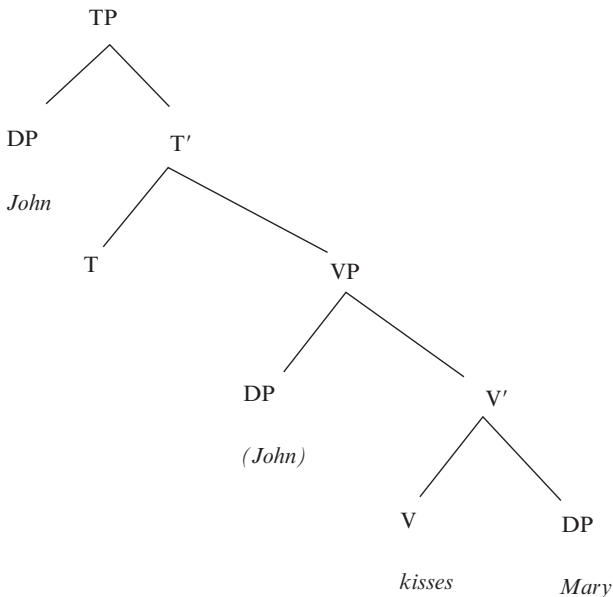
(90)



Here β is contained in γ , the structural sister of α , and so, by (89), α asymmetrically c-commands β .

We can illustrate how the Agree relation works with English subject-verb agreement. The standard relation of subject-verb agreement in English, as in an example like *John kisses Mary*, corresponds to the following structure:

(91)

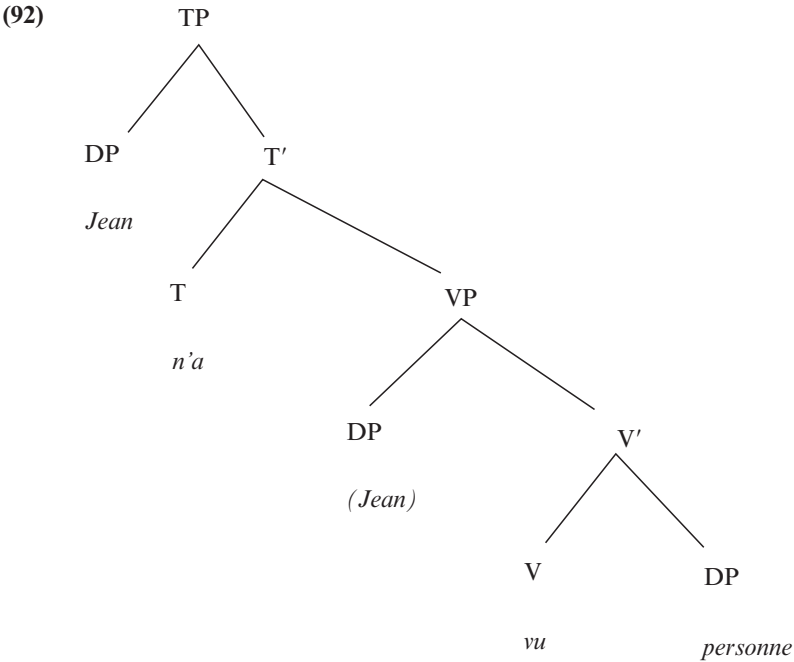


Here, the structural sister of the position occupied by *John* is T' , and the verb is contained in this TP. Hence *John* asymmetrically c-commands the verb, and so the verb's uninterpretable features can be eliminated under Agree with those of *John*. We will see in the next chapter that the Agree relation is more complex than this, being mediated by the T-position; this simplified presentation

suffices to illustrate in general terms how Agree works, however. The Agree relation refers to non-distinctness of features, since *John* is 3rd-person, singular, masculine, while the verb only marks 3rd-person and singular. Features of Person, Number, and Gender are collectively known as φ -features.²³

Where the conditions for Agree as given in (88) are met, α is referred to as the *Probe* and β the *Goal* of the *Agree* relation. Chomsky proposes that Probes and Goals may have uninterpretable features (in addition, possibly, to interpretable features) which render them active for the Agree relation. I will assume here that just the Probe needs to be activated by an uninterpretable feature; a Goal merely requires a non-distinct feature, which may be either interpretable or uninterpretable.

Now, it is natural to suppose that Negation is also a formal feature, and so we can construe negative concord as in the French examples in (87) as an instance of the Agree relation. In all these examples, we take *ne* to occupy T (along with the finite verb – see §1.3.1) and the various negative words *jamais*, *aucun*, and *personne* to be inside VP. Therefore, the c-command condition on Agree is met. This is illustrated in the following structure for (87c):



It is also sensible to regard the negative words (n-words) as having interpretable negative features, since, as mentioned above, they have a

²³ The verb is also present tense, of course. This uninterpretable feature is eliminated under Agree with T.

negative interpretation in isolation. On the other hand, *ne* does not always have a negative interpretation, as for example in sentences like the following:

- (93) Jean est plus intelligent que je ne croyais.
 John is more intelligent than I neg thought
 ‘John is more intelligent than I thought.’

It therefore seems quite reasonable to take *ne* to be the Probe and the n-words to be the Goal in an Agree relation for the feature Negative with *ne* in examples such as those in (87). For convenience, I abbreviate ‘the Agree relation for the feature Negative between α and β ’ as ‘Agree_{Neg}(α , β).’ Agree-based accounts of negative concord are also proposed by Watanabe (2004) and Zeijlstra (2004), both of which are technically more elaborate than what is sketched here.

The idea of *ne* as Probe, Agreeing with a Goal which has an interpretable Negative feature, provides a natural account of the basic pattern of clausal negation in *ne ... pas* in French, as in:

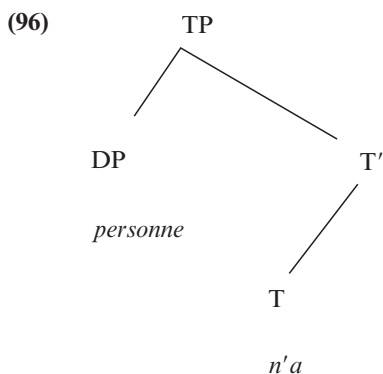
- (94) Je n’ai pas vu Marie.
 I neg-have not seen Marie
 ‘I haven’t seen Marie.’

Here we clearly have Agree_{Neg}(*ne*, *pas*).

However, the Probe-Goal relation appears to be reversed in examples like (95a), and appears to be more complex in (95b):

- (95) a. **Personne** ne m’a vu.
 no-one neg me-has seen
 ‘No-one has seen me.’
 b. **Personne** n’a rien fait.
 no-one neg-has nothing done
 ‘No-one has done anything.’

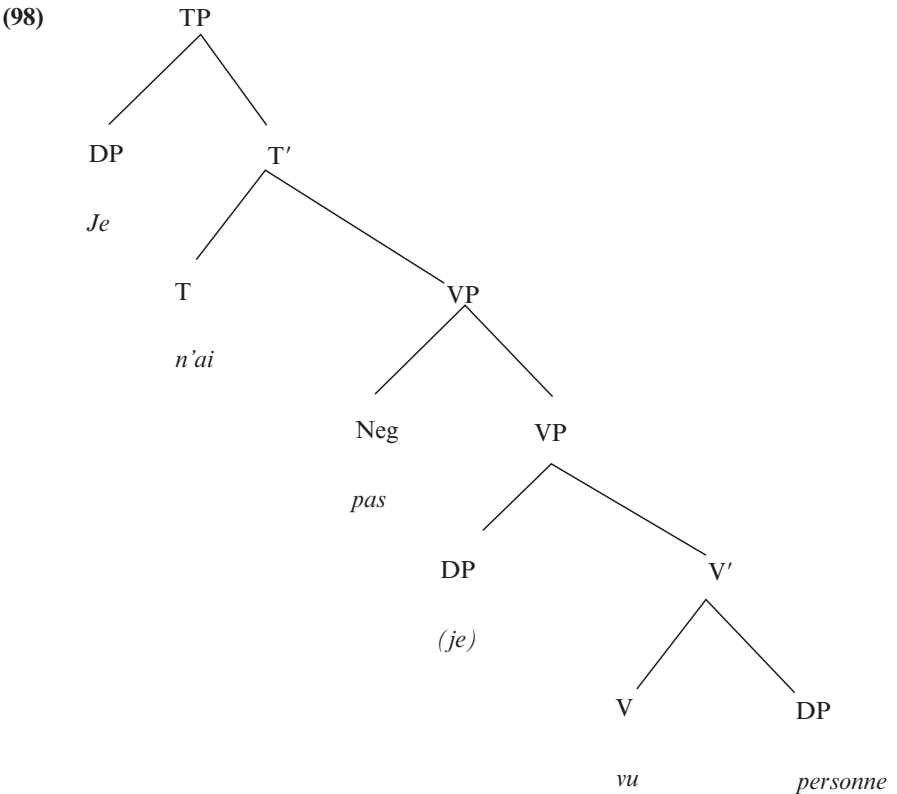
The relevant parts of (95) have the following configuration:



As (96) shows, *personne* c-commands *ne* in (95a). Here, then, we have Agree_{Neg} (*ne*, *personne*), where the uninterpretable Negation feature of *ne* is eliminated, and the interpretable Negation feature of *personne* is retained. In (95b), we have two Agree relations: Agree_{Neg} (*ne*, *personne*) and Agree_{Neg} (*ne*, *rien*). These compose into a single relation Agree_{Neg} ((*ne*, *personne*), (*ne*, *rien*)) expressing a single logical negation. We must assume that in such composed Agree relations, the two interpretable negation features of *personne* and *rien* are identified in such a way as to ensure a single-negation interpretation.²⁴

A third aspect of the Agree relation emerges when we consider examples of the following type:

- (97) **Je n'ai pas vu personne.*
I neg-have not seen no-one



²⁴ Zeijlstra (2004: 247–8) shows how this can follow from a particular proposal for the semantics of n-words. It is also rather similar to what must happen in multiple wh-questions of the type *Who saw what?*, which constitute a single interrogation over two DPs, giving rise to the ‘pair-list’ interpretation we briefly saw in the Introduction to this chapter. See also Box 1.5.

Examples like this are ‘very marginal’ and ‘always have a double negation reading’ (Déprez 1997: 114). The structure of (97) is given in (98).

Here, *ne* c-commands *pas* and *pas* c-commands *personne*. However, these elements are unable to form a single instance of negative concord, unlike in the case of (95b). We can understand this if we take it that *pas*, on the one hand, cannot be a Probe (perhaps because its content is exhausted by its interpretable Negation feature, as opposed to *ne*, which has only an uninterpretable Negation feature), and that, on the other hand, the presence of *pas* blocks the Agree relation between *ne* and *personne*. This last point brings us to the third part of the definition of Agree:

- (88) (iii) there is no γ non-distinct in formal features from α such that γ c-commands β and α c-commands γ .

In the structure in (98) *pas* acts as γ , since it is non-distinct in formal features from α (*ne*), and it c-commands β (*personne*) and is c-commanded by *ne*. In the relevant sense, then, *pas* ‘breaks’ the Agree relation between *ne* and *personne*; this is similar, but not precisely identical to, what Chomsky (2000: 123) calls ‘defective intervention constraints’. On the other hand, Agree_{Neg} (*ne*, *pas*) is allowed, while Agree_{Neg} (*pas*, *personne*) is ruled out perhaps for the reason just given.²⁵

Negative concord thus arises through an Agree relation holding across negative items in a syntactic domain. This is also proposed by Zeijlstra (2004, Chapter 8). But if both the Agree relation and Negation features are invariant aspects of UG, as we surely must assume, then why do languages vary according to the availability of negative concord? Returning to English, we can in fact observe that the Agree relation is at work in certain cases of negative interpretation. The relevant fact is that the expression *anything* in (84b) is a dependent expression. It cannot appear without a special

²⁵ The difference between *rien* and *pas* is important for understanding the grammaticality of (i):

- (i) Il n’a rien dit à personne.
 he neg-has nothing said to no-one
 ‘He said nothing to anyone.’

Here *rien* can act as a Probe for the interpretable negative feature of *personne*, since, unlike *pas*, it has an uninterpretable feature which renders it active. This could be the categorial D- or Q-feature (for quantification), since *rien*, unlike *pas*, has **quantificational force**.

element – which may be a negative, as in (84b) – to ‘license’ it in the same syntactic domain, as the following examples illustrate:

- (99) a. Did you see anyone? (questions)
b. If you see anyone, let me know. (conditionals)
c. John is taller than anyone. (comparatives)
d. *I saw anyone. (not licensed)

Items like *any* are known as **polarity items**, in that they are syntactically and semantically dependent on the presence of another element in the clause which allows them to be interpreted. To go into the precise syntactic and semantic details of how polarity items are licensed would take us too far afield here (for an overview of the relevant literature, see Horn and Kato (2000: 9–11)). However we can observe two things. First, when licensed by negation, the combination of *not* and *any x* gives rise to an interpretation equivalent to that of *no x*. Second, *any* has to be c-commanded by the element which licenses it, as the following examples show:²⁶

- (100) a. *Anyone didn’t see John.
b. Didn’t anyone see John?
c. I didn’t know that anyone had seen John.

It is tempting, therefore, to regard the polarity-licensing relation as a further instance of Agree. However, it seems clear that the licenser has interpretable features, whilst it is the polarity item that has an uninterpretable feature, as shown by the fact that the polarity item cannot stand alone, as in (99d), or act as a Probe for *not* as in (100a). Of course, since a fairly wide class of elements can license polarity items, we have to treat the relevant feature as something more general than Negation; for present purposes, it suffices to call this an uninterpretable Operator feature. See Ladusaw (1980) for a semantic characterization of the class of polarity-licensers in English, and Giannakidou (1998) for a general, cross-linguistic study; in a similar vein, Watanabe (2004: 560) suggests that a Focus feature renders n-words active for Agree in negative-concord languages. We can thus refer to the polarity item as a Probe, and consider the polarity-licensing instance of Agree as an ‘inverse’ Probe–Goal relation. The Agree relation, subject to c-command, matching of non-distinct features and the locality condition of (89iii), holds in the same way in both negative-concord and negative-polarity cases (except for the possibly rather important difference concerning the ‘distance’ of the relation mentioned in Box 1.3). The difference between

²⁶ Hoeksema (2000) takes issue with the c-command generalization for polarity *any*.

BOX 1.3 The interpretation of *any*

It is usually stated that negative-polarity (NPI) *any* is to be distinguished from ‘free-choice’ (FC) *any*. The distinction can be most clearly seen in contexts where there is ambiguity:

- (1) a. If you don’t understand any aspect of the instructions, please let us know.
b. I don’t want to go anywhere.

With NPI *any* (1a) can be interpreted as ‘if you understand *no* aspect of the instructions, ...’, while with FC *any* it means ‘if there is *some* aspect which you don’t understand, ...’. Similarly, (1b), with NPI *any*, means ‘I want to go nowhere’, while with FC *any* it means ‘Not all places are such that I want to go there’, or ‘There are some places I don’t want to go to’. Particularly in (1b), the FC interpretation is much more salient if *any* is stressed. FC also emerges where *any* is modified by *just*, *almost*, or *absolutely*, if *any* modifies a numeral + noun combination (*any three men*), or if it can be followed, colloquially, by *old*. Finally, FC *any* is available in a range of contexts where NPI *any* is not allowed (since it is not c-commanded by an element of the relevant type – see text):

- (2) a. Anyone can play guitar. (‘existential’ modal)
b. Any doctor will tell you. (generic)
c. Ask any doctor. (‘permissive’ imperative)
d. I would give anything for that. (hypothetical)
e. I would have done anything. (counterfactual)
f. Any size fits. (‘sufficient’ conditions)
(This list is taken from Haspelmath (1997: 49ff.).)

On the relation between ‘free-choice’ *any* and ‘polarity’ *any*, see Horn (2000: 157ff.).

The relationship between NPI *any* and its licenser is local, but not clause-bound, as (100c) in the text shows. However, it cannot extend into a subject clause, an adjunct clause, an indirect question, or a relative clause (as pointed out by Ross (1967)):

- (3) a. ?Hiring anyone isn’t allowed.
b. ??John doesn’t want to stay home in order to see anyone.
c. ?I’m not asking you when I should see anyone.
d. I’m not going to sign a petition that any half-baked Stalinist wrote.
(Labov 1972: 144)

(cont.)

BOX 1.3 (cont.)

In each of these cases, to the extent that the sentence is grammatical, *any* must be interpreted as the free-choice variety (as the availability of modification with *absolutely* shows).

This is not true for the relation between *ne* and its associated n-word in French, as the following examples (from Déprez (1997: 111)) illustrate:

- (4) a. *Je ne crois que Marie ait dit qu'elle ait vu personne.
I neg think that Mary have said that she have seen no-one
b. *Je n'ai promis à Jean de rencontrer personne.
I neg have promised to John to meet no-one

In the system presented in Chomsky (2001), Agree is sensitive to 'local' relations of this type, which may further support the idea that negative concord and negative-polarity-item-licensing are instances of Agree (although (100c) remains a problem for this idea). Giannakidou (2000: 469ff.) claimed that the licensing of n-words and the licensing of NPIs are subject to quite distinct locality conditions, with the latter not being clause-bound, while the former is. Zeijlstra (2004: 266–9) discusses the local nature of Agree_{Neg} in Italian.

Not ... any NP and *no NP* may show differences in the scope of negation, as the following contrast shows:

- (5) a. I will force you to marry no-one.
b. I won't force you to marry anyone.

Example (5a) is ambiguous between forced non-marriage and lack of force to marry a particular person, depending on whether negation is construed in the lower or the higher clause. (5b), on the other hand, only has the lack-of-force reading. Compare:

- (6) I will force you not to marry anyone.

Example (6) only has the forced non-marriage reading. The scope of Negation corresponds to the surface position of *not* with *not ... any NP*, but is ambiguous with *no NP*, like many quantifiers.

negative-polarity-item-licensing and negative concord emerges as a difference in the type of Probe–Goal relation that holds involving Negation features, which we can sum up as follows:

- (101) a. English Negative Agree: (α, β) where α is a Goal and β a Probe
b. French Negative Agree: (α, β) where β is a Goal and α a Probe

Example (101) does not, however, state a parametric difference between grammatical systems, as some languages follow neither the negative-concord pattern nor the negative-polarity-item one in the expression of negation, but rather employ negatively quantified expressions comparable to *no x* exclusively. This may be the situation in German, for example (although see Box 1.4). Furthermore, negative-concord languages have polarity items of various kinds; Déprez (1997: 109ff.; 2000: 275ff.) gives examples of negative-polarity items in French.

The parametric property appears to be the existence of French-style Agree_{Neg}. This in turn depends on the existence of a clausal negation which either can or must be uninterpretable, like French *ne*.²⁷ So we can formulate the parameter governing negative concord as follows:

- D. Are (non-inverse) Negative Agree relations found?
 YES: French, Italian, Welsh ...
 NO: English.

BOX 1.4 Cross-linguistic variation in negative concord

We have restricted ourselves to the comparison of English and French, in order to illustrate the basic facts of negative concord and its absence. However, there is considerable variation among negative-concord languages. Italian is rather similar to French, but lacks an equivalent of *pas*. Instead, clausal negation is carried by the preverbal clitic *non*, the equivalent of *ne*. *Non* differs from *ne*, however, in that it cannot co-occur with an n-word in subject position:

- (1) *Nessuno non ha visto Gianni.
 no-one not has seen John

(Compare the French example in (96a)). A possible account for this in terms of the ideas just put forward in the text might claim that *non* optionally has an interpretable Negation feature, and that Agree_{Neg} (*nessuno, non*) is not allowed, as *non* has no content beyond its uninterpretable negation feature. The first idea is needed if we are to allow for negative concord in languages which do not express simple clausal

(cont.)

²⁷ Cf. Zeijlstra (2004: 266), who says ‘if a language has a negative marker that is a syntactic head, the language exhibits N[egative]C[oncord]’.

BOX 1.4 (*cont.*)

negation with two morphemes, i.e. languages like Italian, Spanish, and Greek. The data in (2) below casts some doubt on the second idea, however. Double negation is marginally possible in examples like (1), especially if *nessuno* is stressed. The possibility of double negation is predicted by the Agree_{Neg} analysis put forward in the text.

Many negative-concord languages behave largely like French and Italian but allow an n-word subject to co-occur with the main expression of clausal negation, i.e. they allow the equivalents of Italian (1). This is the case in Rumanian (as well as Hungarian, Greek, and many Slavonic languages), shown in (2):

- (2) Nimeni* (nu) a venit a petrecere.
 no-one not has come to the-party
 ‘No-one came to the party.’
 (Martins (2000: 196))

It is unclear what underlies the difference between such languages and Italian, as just described.

Some authors (Giannakidou (1997; 2000); Zeijlstra (2004)) refer to Italian-type languages as ‘non-strict negative-concord’ languages and to Rumanian-type languages as ‘strict negative-concord’ languages. Zeijlstra (2004: 244–61) proposed an analysis of the difference between strict and non-strict negative concord using Agree, but making rather different technical assumptions from those adopted here.

It is also possible to distinguish ‘negative spread’ from ‘negative concord’. Negative spread involves co-occurrence of two negative expressions with a single negative interpretation, in the absence of a clausal negation marker. This is found in German (while ‘standard’ negative concord is not):

- (3) Hier hilft KEINER KEINEM
 here helps no-one no-one
 ‘Nobody helps anybody here.’

(The capital letters indicate that the relevant parts of the sentence must be pronounced with heavy stress.) I will not speculate here as to how to analyse this phenomenon using the mechanism of Agree as described in the text. On negation in Welsh, see Borsley and Morris-Jones (2005); on its historical development, see Willis (forthcoming).

In this section we have seen two main things: first, the concept of negative concord and the observation that this is a property of some languages but not others, and second, the technical notion of Agree and the associated distinction between interpretable and uninterpretable features as a way of analysing negative concord (and possibly negative-polarity-item-licensing as well). This led us to the postulation of parameter D, to whose diachronic effects we now turn.

1.4.2. *Negative concord in the diachronic dimension: the development of French n-words*

Here I want to concentrate on the history of French. It appears that the expressions which are n-words in Modern French (*rien*, *personne*, *jamais*, etc., as discussed in the previous section) were, at an earlier stage in the history of the language, either indefinites or negative-polarity items. The fact that these words are not historically negative can be seen from their etymologies: *rien* comes from Latin *rem*, meaning ‘thing’; *personne* is etymologically related to the feminine noun *personne* (‘person’ – the English word is of course borrowed from French); and *jamais* derives from a compound of *ja* (‘already’) and *mais* (‘more’). The other n-words of French have similar non-negative etymologies, with the single exception of *nul* ‘no’ from Latin *nullus* ‘no’.

Regarding the nature of these elements, Foulet (1990: 244) comments as follows:

Although *ne* is the essential negation in Old French and needs no extra help to express the idea of negation, it is nevertheless the case that from an early stage there is a preference to reinforce it with a series of words whose usage is sometimes rather curious. These words, with one exception, . . . take their negative value purely from their association with *ne*, and it is impossible to use them with a negative meaning without *ne* preceding or following them.²⁸

²⁸ ‘Si *ne* est la négation essentielle du vieux français et n’a besoin d’aucun secours étranger pour exprimer l’idée négative, il est vrai pourtant que depuis longtemps on aime à la renforcer par une série de mots dont l’emploi est parfois bien curieux. Ces mots, à une exception près, . . . tiennent leur valeur négative uniquement de leur association avec *ne*, et il est impossible de les employer au sens négatif sans les faire précéder ou suivre de *ne*’ [my translation].

In stating that *ne* is the ‘essential negation’ of Old French, Foulet means that it alone sufficed to express clausal negation, unlike in Modern French. This is illustrated by examples like (102):

- (102) a. Je **ne** nourriroie trahitor.
 ‘I would not feed [a] traitor.’
 (Ch. 1223–4; Foulet 1990: 73)
- b. Li ostes **ne** set que il vent.
 the landlord not knows what he sells
 ‘The landlord doesn’t know what he is selling.’
 (Bodel *Le jeu de saint Nicholas*, l. 700; Ayres-Bennett 1996: 72)
- c. Un moyne ... **ne** laboure comme le paisant, **ne** garde le pays
 a monk not works like the peasant, not guards the country
 comme l’homme de guerre, **ne** guerit les malades
 like the soldier not cures the sick
 comme le medicin, **ne** presche ny endoctrine le monde comme le
 like the doctor, not preaches nor teaches the world like the
 bon docteur evangelicque et pedagogue, **ne** porte les commoditez
 good doctor evangelical and pedagogical, not brings the commodities
 et choses necessaires à la republicque comme le marchand.
 and things necessary to the republic like the merchant
 ‘A monk ... doesn’t work like the peasant, doesn’t protect the land like
 the soldier, doesn’t cure the sick like the doctor, doesn’t preach or
 teach like the good preacher and teacher, doesn’t bring goods and
 commodities essential to the nation like the merchant.’
 (Rabelais *Gargantua* (ed. Calder, Droz 1970: 229); Ayres-Bennett
 1996: 143, her translation)

From this we must conclude that *ne* had an interpretable Negation feature at this period. The example in (102c), which is from the sixteenth-century author Rabelais, shows that this was the case at least until this period. In fact, it is generally stated that *ne* ... *pas* negation became obligatory in the seventeenth century (Ayres-Bennett 1996: 146; Harris 1978: 26; Robert 1992: 1441). (See note 7, Chapter 3.)

In the quotation given above, Foulet also points out that the words which ‘reinforce’ negation are not negative unless *ne* appears in the same context. This can be seen from examples such as the following:

- (103) a. comment qu’il **onques** en aviegne
 how that it ever of-it happens
 ‘how it might ever happen’
 (*Courtois d’Arras* 66; Foulet, 252)

- b. **Aucuns** se sont aati ...
 some selves are boasted ...
 ‘Some (people) have boasted ...’
 (le Bossu, *Jeu de la Feuillée* 438; Foulet, 246)
- c. douce **riens** por cui je chant
 sweet thing for whom I sing
 ‘sweet one for whom I sing’
 (Muset, *Chansons* VIII, 44; Foulet, 273)

In (103a), *onques* is either an indefinite adverb or a polarity item; since it appears in an indirect question, we cannot tell which it is on the basis of this example. In (103b), it is fairly clear that *aucuns* is an indefinite, while in (103c) *rien* is a noun (feminine in gender, as the form of the adjective *douce* shows). These examples clearly show that these items did not at this time have an interpretable Negation feature, although, if some of them were polarity items, they may have had an uninterpretable Operator feature, as was proposed for English *any* in the previous section.

These expressions appear in negative clauses in OF where *ne* is present:

- (104) a. Ce n'avint **onques**.
 that not-happened ever
 ‘That didn’t ever happen/that never happened.’
 (*Chastelaine de Vergi* 349; Foulet, 252)
- b. k'il n'aient de vous **aucun bien**
 that-they not-have from you a good
 ‘that they won’t have any good(s) from you’
 (*Jeu de la Feuillée* 671; Foulet, 247)
- c. ... li feus, qu'il **ne** pooit por **riens** estaindre
 ... the fire that-he not could for thing put-out
 ‘... the fire that he couldn’t put out for anything’
 (Huon le Roi, *Le Vair Palefroi* 204–5; Foulet 279)

As the English translations in *ever* and *any* indicate, it may be possible to consider *onques*, *aucun*, and *rien* as polarity items in these examples. Alternatively, they may be regarded as indefinites. In fact, negative-polarity items may themselves be a kind of indefinite; see Giannakidou (2000) and Horn (2000) for a critical discussion of this question. But, given the evidence in (103), they cannot be n-words in the way they are in Modern French.

It seems, then, that in OF *ne* was the true negation, and thus had an interpretable Negation feature. The fact that the future n-words could

appear in non-negative contexts like (103) shows that they did not at this time have an interpretable Negation feature. If they, or some of them at least, were polarity items, then they may have had an uninterpretable Operator feature, as already mentioned. What seems to have happened is that in examples like (104) the negative force was – loosely speaking – ‘transmitted’ to the former indefinites from *ne*, which thus became n-words.

We can make sense of this notion of the ‘transmission’ of negative force from *ne* to the n-words in terms of the analysis of negative concord and negative-polarity using Agree which we put forward in the previous section. What happened was that the interpretable Negation feature became associated with the n-words, and dissociated from *ne*. Arguably the crucial aspect of this change is that *ne*’s Negation feature became uninterpretable. Because of this, French developed the Negative Agree relation which constitutes a positive setting for parameter D. So we see that parameter D has changed its value in the history of French.

If the crucial step in this change is the development of an uninterpretable negative feature on *ne*, we might expect the development of negative concord (and hence n-word status for *aucun*, etc.) to correlate chronologically with the development of the *ne ... pas* clausal negation. This appears to be true. As mentioned above, the *ne ... pas* negation became obligatory in the seventeenth century; *ne* had lost its interpretable Negation feature by this time. It is possible that *ne* passed through a stage in which this feature was optionally interpretable (as suggested for Italian in Box 1.4), but this cannot be determined with any certainty on the basis of the available data.

Martins (2000: 195–8) shows that other Old Romance languages (Old Spanish, Portuguese, Galician-Leonese, and Italian) were like OF in allowing what are now n-words to appear in ‘modal contexts’, i.e. ‘questions, imperatives, conditionals, comparatives, the scope of modal verbs, the scope of words expressing prohibition, generic constructions, subjunctive clauses introduced by ... “before”’ (195). It may be, then, that these elements were polarity items at this period. In terms of the analysis presented in the previous section, this would imply that they lacked a Negation feature but had an uninterpretable Operator feature at the earlier period.

Martins also shows that Modern Catalan has this pattern. Interestingly, some of the polarity items in Catalan have a non-negative etymology, like those in French. This is true of *res* (‘nothing’), like French *rien* from Latin *rem*, and *cap* (‘no’), as in:

- (105) a. T'ha passat res?
 to-you-has happened nothing
 'Did anything happen to you?'
 b. Si hi trobeu cap defecte, digueu-m'ho.
 if in-it you-find no defect, tell-me-about-it
 'If you find any defect, let me know.'

These elements cannot appear in (non-modal) positive declarative clauses:

- (106) *T'ha passat res.
 to-you-has happened nothing

Instead, the unambiguously positive *alguna cosa* ('something') must be used here. This situation appears to resemble OF, as Martins points out.

It appears, then, that parameter D has changed its value in the history of French, probably around 1600. It may also have changed its value in the history of other Romance languages, with the exception of Catalan. In French, the change in this parameter is connected to the development of the *ne ... pas* negation. The development of this negation pattern, and the subsequent loss of *ne* in Colloquial French, is an instance of Jespersen's Cycle (see Jespersen (1917)); I will discuss this in §2.2.

As I mentioned at the beginning of this section, parameter D is rather different from the others discussed here, as it does not directly concern word order, but rather variation in the feature-content of certain classes of lexical items. It is also more closely connected with semantics than the other parameters we have looked at.

1.5. Wh-movement

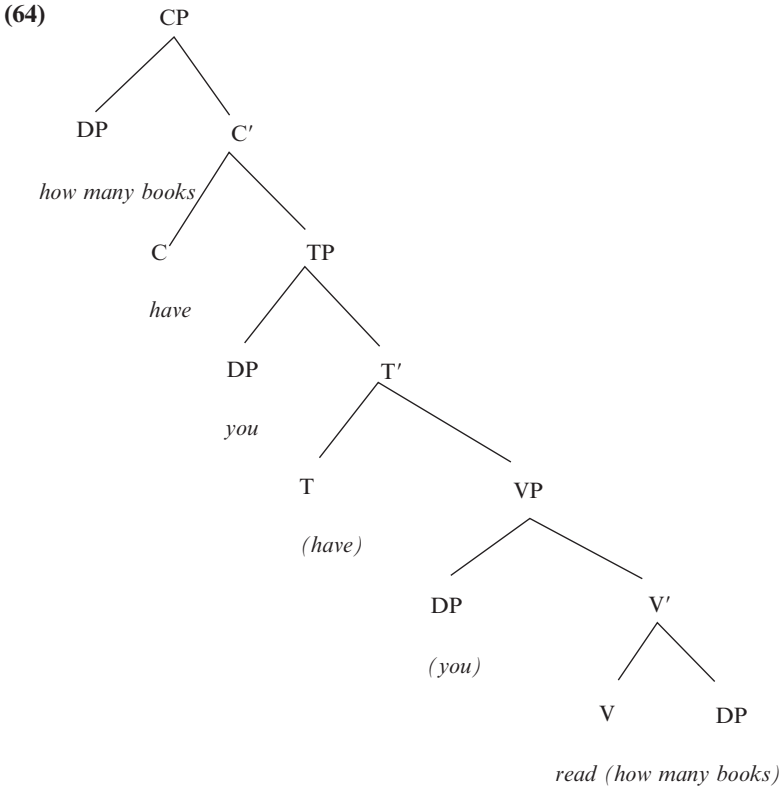
1.5.1. The *wh*-movement parameter

An important parameter, which has attracted a great deal of attention since it was first proposed in Huang (1982), concerns the incidence of 'overt', as opposed to 'covert' *wh*-movement.

Wh-movement is an operation that takes place in, among other constructions, *wh*-questions. We came across this operation in our discussion of the null-subject parameter in §1.1.1 and again in our discussion of residual V2 in §1.3.1.2. I repeat (63a) here for convenience:

(63) a. How many books have you read?

As we saw, (63a) has the structure in (64), also repeated here:



In this structure, movement of the DP *how many books* to SpecCP is an instance of wh-movement, leaving a copy in the direct-object position, as shown here. Wh-movement can also take place in indirect questions, although here there is no T-to-C movement in Standard English – see §5.2.3 on some non-standard varieties:

(107) I wonder [_{CP} how many books he has read (how many books)]

In English, movement of one wh-phrase to SpecCP is obligatory in both direct and indirect wh-questions, as the ungrammaticality of (108) shows:

(108) *Have you read how many books?

If we ‘undo’ T-to-C movement, we have a well-formed ‘echo question,’ as distinct from a simple request for information:

(109) You have read HOW many books?!

As the punctuation is meant to indicate, (109) is only grammatical with a very particular intonation. Indirect echo questions are impossible:

(110) *I wonder he has read HOW many books?!

In the clausal complement to verbs like *wonder*, wh-movement to SpecCP is required.

Many languages do not have wh-movement of the kind we find in English, although they certainly allow their speakers to ask wh-questions. In these languages, the wh-phrase remains unmoved, in its grammatical-function position (subject, direct object, adverb, etc.). Chinese and Japanese are perhaps the best known examples of such languages, called *wh-in-situ* languages, as illustrated by the examples of indirect questions in (113):

- (111) a. Zhangsan xiang-zhidao [Lisi mai-le sheme]. (Chinese)
 Zhangsan wonder Lisi bought what
 ‘Zhangsan wonders what Lisi bought.’
 (Watanabe 2001: 203, (2))
- b. Boku-wa [_{CP} [_{IP} John-ga nani-o katta] ka] shiritai. (Japanese).
 I-Top John-Nom what-Acc bought Q want-to-know
 ‘I want to know what John bought.’
 (Watanabe 2001: 207, (15a))

We can thus formulate a parameter to distinguish English-style languages from Chinese-Japanese style languages:

E: Does a wh-phrase move to the Specifier of an interrogative CP?

YES: English, Italian, Spanish, German, Welsh ...

NO: Chinese, Japanese, Thai, Korean, Turkish ...

What properties co-occur with parameter E? One possibility is that *in-situ* languages lack wh-determiners and wh-pronouns of the kind exemplified by *which*, *what*, etc. in English. The Chinese and Japanese words that translate these English expressions are in fact indefinite pronouns; the interpretation of the clause as a wh-question depends on the presence and nature of a sentential particle in many languages. This is clearly illustrated in Chinese, where the nature of the particle determines the interpretation of the clause, and of the indefinite DP, as involving wh-interrogation or not:

(112) a. Hufei chi-le sheme (ne)

BOX 1.5 Cross-linguistic variation in overt wh-movement

In English, it is possible for more than one wh-phrase to appear in a single sentence. We saw an example of this in the Introduction to this chapter, which I repeat here, with the copy of wh-movement, *who*, indicated:

(1) Who did he think (who) was likely to drink what?

As we saw, examples like these can have a ‘pair-list’ interpretation. What is impossible, however, is movement of more than one wh-phrase to SpecCP:

(2) *Who what did he think (who) was likely to drink (what)?

But some languages allow this. The best known ‘multiple wh-movement’ languages are the Slavonic languages, illustrated by the following Bulgarian example (from Rudin (1988)):

(3) Koj kogo e vidjal?
 who whom aux saw-3s
 ‘Who saw whom?’

Here a further parameter must be at work, distinguishing among languages with the positive value for parameter E (but see Bošković (2002) for a different view). Since I am not aware of a clear example of diachronic change between an English-style ‘single wh-movement’ and a Slavonic-style multiple wh-movement system I will say no more about this option here.

Other parameters clearly distinguish among languages with wh-movement. One, which we will look at in more detail in §3.3, concerns the possibility of Preposition-stranding. Preposition-stranding is the cross-linguistically rather rare option of moving the complement of a preposition, while leaving the preposition ‘stranded’. English and Mainland Scandinavian languages allow this in wh-questions, illustrated in (4a) (it is also allowed in passives, but I will leave that point aside here):

(4) a. Who did you speak to ___?
 b. To whom did you speak ___?

As (4b) shows, English also allows ‘**pied-piping**’ of the Preposition along with the wh-phrase. It is not clear what exactly permits this in English and the Scandinavian languages. In most other languages which have

overt wh-movement, the preposition must be pied-piped. This is the case of Standard French, for example:

- (5) a. *Qui as-tu parlé à ___?
 b. A qui as-tu parlé ___?

French has some property which requires pied-piping of the DP in all cases. Kayne (1984) made a very interesting and influential proposal concerning this, but in the context of technical assumptions which have not been carried over into most versions of minimalism. We will look at Preposition-stranding in one variety of French in our discussion of contact in §3.3.

Another parameter concerns the possibility of ‘left-branch extraction,’ i.e. moving a wh-phrase from the left branch of the category containing it. This is not allowed in English, as (6) shows:

- (6) *Whose did you read [_{DP} (whose) book]?

Here *whose* has moved from the left branch of the DP. The equivalent of (6) is allowed in many languages, however. (7) is an example from Russian:

- (7) Č’ju ty čital knigu?
 whose you read book
 ‘Whose book did you read?’
 (Biberauer and Richards (2006: 21, (28)))

We will encounter this fact about Russian again in §4.1.4.

- Hufei eat-asp what Q_{wh}
 ‘What did Hufei eat?’
 b. Qiafong mai-le *sheme* **ma**
 Qiafong buy-asp what $Q_{y/n}$
 ‘Did Qiafong buy anything?’
 (Cheng (1991: 112–13))

Here we see that the choice of the sentence-final particle determines whether *sheme* is interpreted as an indefinite in a yes/no question or as a wh-phrase in a wh-question. Cheng (1991) makes two interesting generalizations in this connection. First, she proposes that every clause needs to be ‘typed’, i.e. its ‘force’ as an interrogative, declarative, etc. must somehow be marked. In the case of typing a wh-question, she proposes that either a wh-particle in C or fronting of a wh-word to the SpecCP is used (30). Assuming that C Agrees for a wh-feature with a given DP, we could think

of this as realization of the *wh*-feature on either the Probe or the Goal, with associated movement of the Goal when the *wh*-feature is morphologically realized there. This implies that C is final in the Chinese examples in (112) above; we will turn to the question of cross-linguistic variation in word order in the next section.

Another potentially important property connected with *wh*-movement is word order in the VP. Bach (1971) observed that *wh-in-situ* languages tend to be OV (on OV languages in general, see §1.6). He formulated what has become known as Bach's generalization:

(113) If a language is OV, then it has *wh in situ* (see Kayne (1994: 54)).

This is an example of a possible implicational universal, something I will come back to in much more detail in §1.6.

Languages like Dutch, German, and Latin are problematic for (113) since they clearly have *wh*-movement and are OV (see §1.6.1). Haspelmath *et al.* (2005) provide data on the correlation between VO order and overt *wh*-movement. Of the 711 languages with determinate properties regarding both *wh*-movement and OV or VO word order, they observed the following pattern of co-occurrence:

(114)	OV and <i>wh in situ</i>	280
	VO and <i>wh in situ</i>	219
	VO and <i>wh</i> -movement	143
	OV and <i>wh</i> -movement	69

Most of the languages combining OV and *wh*-movement are Australian or Amerindian (along with some Khoisan and Nilo-Saharan languages), which may explain why this possibility was earlier thought to be rare or non-existent; these languages have become much better analysed in recent years. We can observe, however, that it is less common than the other patterns, with less than 10 per cent of the surveyed languages showing it. Bach's generalization, then, may indicate a tendency in the world's languages. As we will see in the next section, Watanabe (2002) suggests Old Japanese did not obey this generalization.

Finally, it has been suggested that 'wh-agreement' is only found in contexts of *wh*-movement, and never in *wh-in-situ* languages. Wh-agreement can be defined as 'a phenomenon in which verbal inflection and complementizers display distinct morphosyntactic properties in the clauses which immediately contain the displaced *wh*-phrase and its traces [i.e.

copies – IGR]’ (Watanabe 2002: 183). A familiar example is past-participle agreement in compound tenses in French, which, at least in literary French, is found when a direct object undergoes *wh*-movement (as well as in other contexts, which I leave aside here):

- (115) *Quelle voiture as-tu conduite?*
 which car (f.sg.) have-you driven (f.sg.)
 ‘Which car have you driven?’

In Irish, the complementizer introducing a finite clause from which *wh*-movement has taken place changes its form from *go* to *aL* (the ‘L’ indicates that this element causes lenition in the initial consonant of the following word – see McCloskey (2001) for details):

- (116) a. *Deir said gur ghoid na síogaí í.*
 say they C-[PAST] stole the fairies her
 ‘They say that the fairies stole her away.’
 b. *an ghirseach a ghoid na síogaí*
 the girl aL stole the fairies
 ‘the girl that the fairies stole away’
 c. *rud a gheall tú a dhéanfá*
 thing aL promised you aL do [COND-S2]
 ‘something that you promised that you would do’
 (McCloskey 2001: 67–8)

In Palauan, when any argument other than the local subject (in a direct question, the main-clause subject) undergoes *wh*-movement, the verb shows irrealis morphology. This is illustrated in (117a), where the direct object is *wh*-moved, and in (117b), where the embedded subject is moved (here the main verb shows irrealis marking):

- (117) a. *ng-ngerá [a le- silseb-ii (ng-ngerá) a se’el-il]*
 CL-what IR3-PF-burn-3sg friend-3sg
 ‘What did his friend burn?’
 (Georgopoulos (1991: 70), in Watanabe (1996: 171))
 b. *ng-te’a [a l-ilsa a Miriam [el milnguui er a buk er ngii (ng-te’a)]]*
 CL who IR3-PF-see Miriam C R-1M-read P book her
 ‘Who did Miriam see reading her book?’
 (Georgopoulos (1991: 90), in Watanabe (1996: 172))

Watanabe (2002: 184) concludes that ‘if a language exhibits a *wh*-agreement phenomenon, it always shows sensitivity to overt *wh*-movement’. We

will return to this point when we take up Watanabe's discussion of Old Japanese in the next subsection.

In this section we have seen the basic cross-linguistic properties of wh-movement: the existence of wh-movement and wh-*in-situ* languages, a difference captured by Parameter E; a further distinction among wh-movement languages between 'single' and 'multiple' wh-movement; Bach's generalization and Watanabe's observation that wh-agreement may correlate with wh-movement. There is much more to say about wh-movement; in fact the investigation of this phenomenon has been a central theme in generative grammar for many years (see Box 1.6), but the above points are sufficient here.

BOX 1.6 The significance of wh-movement

Although not directly relevant to the aims of this book, it may be worth briefly indicating why wh-movement has been of such importance for generative theory. The basic reason is that wh-movement appears to be able to operate over an unlimited amount of material. For this reason, it is known as an **unbounded dependency**. The apparently unbounded nature of wh-movement is illustrated in (1):

- (1) a. What did Bill buy (what)?
 b. What did you force Bill to buy (what)?
 c. What did Harry say you had forced Bill to buy (what)?
 d. What was it obvious that Harry said you had forced Bill to buy (what)?

However, beginning with Ross (1967), a range of constructions in which even unbounded dependencies cannot be formed was recognized. These constructions are known as **islands**. Two examples of the various islands that were first recognized in Ross (1967) are given in (2) and (3):

- (2) *The complex DP constraint:*
 a. *Which writer did you write [_{DP} a play which [_{TP} was about (which writer)]] ?
 b. *Which writer did you believe [_{DP} the claim that [_{TP} we had met (which writer)]] ?
- (3) *The left branch condition (LBC):*
 a. *Whose did you play [_{DP} (whose) guitar]?
 b. Mick's friend's favourite guitar
 c. Whose guitar did you play (whose guitar)?

Ross (1967) also suggested that the locality condition on the *not ... any NP* relation mentioned in Box 1.4 might be a further case of an island

constraint. ‘Island effects’ of this type are not found in *wh-in-situ* languages such as Chinese, as Huang (1982) first showed. However, very intriguingly, Huang also showed that other locality effects associated with *wh*-movement are found in *wh-in-situ* languages. In particular, adjunct *wh*-elements cannot be interpreted with wide scope in certain islands. The following example illustrates this for the adjunct *weishenme* (‘why’) in a complex NP in Chinese:

- (4) Ni zui xihuan [weishenme mai shu de ren]?
 you most like why buy book Prt person
 ‘Why do you like the man who bought the books?’

This demonstrates that, while movement is sensitive to islands, there are further locality constraints on *wh*-interpretation which are independent of overt movement. See Watanabe (2001) for more systematic discussion of this.

Chomsky (1973) proposed an account of island constraints in terms of the **subjacency condition**, a general condition on the formation of unbounded dependencies. The idea is that, despite appearances, *wh*-movement does not in fact apply over unbounded distances, but instead proceeds in a series of short ‘hops’ via each SpecCP position. Thus an example like (1c) above has the structure in (5), with the various copies indicated:

- (5) [_{CP1} What [_{C1} did [_{TP1} Harry say [_{CP2} (what) [_{C2} [_{TP2} you had forced
 Bill [_{CP3} (what) [_{C3} [_{TP3} to buy (what)]]]]]] ?

This kind of movement is known as successive-cyclic movement. Subjacency is then formulated so as to prevent movement of a *wh*-constituent across more than one blocking category (BC). The blocking categories (in English) are DP and TP. Movement from a complex DP as in (2) crosses a DP and a TP boundary, as does movement from the left branch of a DP as in (3), hence these examples are ruled out by the subjacency condition. Chinese examples like (4) cannot be subject to subjacency, and so some other principle must be at work here (see the textbooks on government-and-binding theory cited in the introduction for discussion of this). It has been claimed that phenomena such as the change in the Irish complementizer illustrated in (116) are evidence for successive-cyclic movement, since exactly the complementizers in clauses containing the position from which the *wh*-constituent was moved show the complementizer change.

In general, *wh*-movement has been important as it has provided clear evidence for the local nature of syntactic dependencies, even where they appear to hold over unlimited domains. It has also provided a clear indication of the precise nature of the locality constraints that hold. It has thus been a central tool in the investigation of UG.

1.5.2. *Wh-movement in the diachronic domain: Old Japanese*

In a recent paper, Watanabe (2002) argues that parameter E changed its value in Old Japanese, specifically between the Nara Period (eighth century) and the Heian Period (ninth to tenth centuries). In the earlier period, then, Japanese had overt *wh*-movement. The basic evidence for this is the observation, which Watanabe attributes to Nomura (1993), that ‘the *wh*-phrase must precede the nominative subject in the Nara Period’ (2002: 181). This is illustrated by examples such as the following:

- (118) a. Kasugano-no fuji-ha chiri-ni-te nani-wo-ka-mo
 Kasugano-GEN wisteria-TOP fall-PERF.CONJ what-ACC-KA-MO
 mikari-no hito-no ori-te kazasa-mu?
 hiker-GEN person-NOM pick-CONJ wear.on.the.hair-will
 ‘Since the wisteria flowers at Kasugano are gone, what should hikers
 pick and wear on their hair?’
- b. Kado tate-te to-mo sashi-taru-wo izuku-yu-ka imo-
 gate close-CONJ door-also shut-PAST.ACC where-through-KA wife-
 ga iriki-te yume-ni mie-tsuru?
 NOM enter-CONJ dream-LOC appear-PERF
 ‘From where did my wife come and appear in my dream, despite the fact
 that I closed the gate and shut the door?’
 (Man’youshuu nos. 1974 and 3117; Watanabe 2002: 182, (5a, b))

The *wh*-phrase here bears the focus marker *ka*. In fact, focused non-*wh*-phrases, marked with *ka*, occupy this pre-subject position, which also follows a topic marked with *hu*:

- (119) Hatsuse-no kawa-ha ura na-mi-ka fune-no
 Hatsuse-GEN river-TOP shore absent-ness-KA boat-NOM
 yori-ko-nu?
 approach-come-NEG
 ‘Is it because Hatsuse River has no shore that no boat comes near?’
 (Man’youshuu no. 3225; Watanabe 2002: 183, (7b))

The existence of overt *wh*-movement in Old Japanese explains the presence of a *wh*-agreement phenomenon, known in the Japanese philological tradition as *kakasimusubi*. *Kakasimusubi* relates the nature of a preposed XP to the form of the verb: where the preposed XP is marked with *ka*, for example, the verb must appear in the ‘adnominal form’, which is

morphologically distinct from the ‘conclusive form’ of (unfocused) declaratives. Watanabe (2002: 180) gives further details. Both the particles and the special verb forms were lost by the fifteenth century, although Watanabe argues that the wh-movement system triggering *kakasimusubi* was lost by the beginning of the eleventh century (181). As he points out, ‘[r]ecognising overt wh-/focus-movement in Old Japanese makes it possible to understand why it used to have the system of *kakasimusubi*’ (183); the wh-agreement was lost as a consequence of the loss of overt wh-movement, i.e. the change in parameter E.

Watanabe documents five changes which take place between the Nara and Heian Periods:

- (120) a. loss of overt wh/focus-movement;
 b. decrease in the use of *ka* with the wh-phrase in genuine wh-questions;
 c. use of wh-*ka* in rhetorical questions;
 d. loss of *ka* with non-wh forms;
 e. increase in subject topicalization.

Examples (120b–d) amounted to the loss of *ka* as the morphological trigger for wh-/focus-movement. (120e) removed the word-order trigger for focus/wh-movement, since it created orders in which the focused XP did not systematically precede the subject. The changes in (120) combined with the consistent verb-final order (which, as we saw in our discussion of Bach’s generalization in the previous section, disfavors overt wh-movement) to eliminate overt wh-movement from the grammar of Japanese. In this way, the value of parameter E changed.

Wh-movement has undergone, or may be undergoing, change in both French and (Brazilian) Portuguese. These are, at least normatively, wh-movement languages like English, showing movement of exactly one wh-phrase per wh-interrogative. However, they both seem to have developed wh-*in-situ* questions in their fairly recent history. Foulet (1921) discusses this in relation to French and Rossi (1993: 328ff.) in relation to Brazilian Portuguese. However, in both languages wh-*in-situ* is restricted to direct questions, i.e. to main clauses, and so the system is not the same as that found in languages such as Japanese, where wh-phrases remain *in situ* in all types of clauses.

Mathieu and Sitaridou (2005) show that Classical Greek allowed left-branch extraction, as in (121b) (see Box 1.5 on this), while Modern Greek does not, as (122) shows:

- (121) a. Tina dynamin echei?
what.ACC.FEM.SG power.ACC.FEM.SG have.3SG
'What power does it have?'
(Plato, *Laws*, 643a; Mathieu and Sitaridou 2005: 237, (1a))
- b. Tina echei (tina) dynamin.
(Plato, *Republic*, 358b; Mathieu and Sitaridou 2005: 237, (1b))
- (122) a. Ti dinami exi?
what power.ACC.FEM.SG have.3sg
'what power does he/she/it have?'
- b. *Ti exi (ti) dinami?
(Mathieu and Sitaridou 2005: 238, (2))

Mathieu and Sitaridou relate the change to the loss of rich agreement on wh-elements in Greek and the reanalysis of *ti*, the definite article in Modern Greek, from an Adjective to a Determiner, which had the consequence that *ti* lost its earlier 'indefinite' uses. (See Roberts and Roussou (2003: 161–7) for details.) We return to the change discussed by Mathieu and Sitaridou in §4.1.4.

We conclude that parameters connected to wh-movement, most importantly parameter E, are subject to diachronic change. Watanabe's discussion of the development of Japanese seems to show this quite clearly.

1.6. Head-complement order

As the final example of parametric variation, I take one of the most pervasive and well-studied instances of cross-linguistic variation: the variation in the linear order of heads and complements. I will concentrate on the aspect of this variation which has been recognized since W. Lehmann (1973) as the most important: the relative order of verbs and their objects. In fact, we took the variation between verb-object (VO) languages and object-verb (OV) languages as our initial example of a parameter in §1.1. Here I develop the points made there in more detail.

1.6.1. Head-complement order synchronically

As we saw in §1.1, English and German differ regarding the order of infinitival verbs and their direct objects. This difference is illustrated by the contrasts in (6) and (7), repeated here:

- (6) a. Tomorrow John will visit Mary.
 b. *Morgen Johann wird besuchen Maria.
- (7) a. Morgen wird Johann Maria besuchen.
 b. *Tomorrow will John Mary visit.

These examples show us that German has OV order in infinitives, while English has VO. Since we have seen that the verb moves to C in main finite clauses in German (see §1.3.1), giving rise to verb-second order, we can suppose that the verb occupies its unmoved position in examples like (7a). Finite verbs may also occupy this position in clauses where movement to C is not allowed. Again, we saw this in §1.3.1 in examples such as (59) and (60):

- (59) Du weißt wohl,
 You know well
 a. ... daß ich schon letztes Jahr diesen Roman **las**.
 ... that I already last year this novel read
 b. ... daß ich schon letztes Jahr diesen Roman **gelesen habe**.
 ... that I already last year this book read have
- (60) Ich frage mich,
 I ask myself
 a. ... ob ich schon letztes Jahr diesen Roman **las**.
 ... if I already last year this book read
 b. ... ob ich schon letztes Jahr diesen Roman **gelesen habe**.
 ... if I already last year this book read have

Indeed, in our representation of a verb-second clause in (62) we assumed that the finite verb was merged in a position following the direct object.

German thus combines OV and V2 orders. These two properties are nevertheless independent. We saw that the Scandinavian languages are V2 but VO, and many OV languages are not V2. Examples of OV languages in which all clauses – main and subordinate – are OV include Japanese, Korean, and Turkish, as sentences like (123) show:

- (123) a. Sensei-wa **Taro-o** **sikata**. (Japanese)
 teacher-TOP Taro-ACC scolded
 ‘The teacher scolded Taro.’
- b. Kiho-ka **saca-li-l** **cha-ass-ta**. (Korean)
 keeho-NOM lion-ACC kick-PAST-INDIC
 ‘Keeho kicked the/a lion.’
- c. Ahmet **kitab-i** **oku-du**. (Turkish)
 Ahmet book-ACC read-PAST
 ‘Ahmet read the book.’

We can thus formulate the OV/VO parameter as follows:

- F1. Do direct objects precede or follow their verbs in overt order?
PRECEDE: German, Dutch, Japanese, Korean, Turkish, Basque ...
FOLLOW: English, Romance, Thai, Zapotec ...

However, as we have already seen, other word-order properties are correlated with the OV/VO parameter. We saw in §1.1 that the relative order of auxiliaries and their associated verbs correlates with OV or VO order, in that VAux order is characteristic of OV languages and AuxV order of VO languages. Controlling once more for the effects of verb-second order in finite main clauses, we can observe this correlation by comparing English and German. The following examples, again repeated from §1.1, illustrate:

- (8) a. John **can visit** Mary. (AuxV)
b. Johann wird Maria **besuchen können**. (VAux)

Similarly, Japanese has VAux order, as in:

- (124) John-ga Mary-to renaisite iru. (VAux)
John-NOM Mary-with in-love is
'John is in love with Mary.'

Other languages showing the combination of OV and VAux orders include Basque, Burushaski, Chibcha, Hindi, Kannada, Nubian, Quechua, and Turkish. These are languages from Greenberg's (1963) original 30-language sample, as summarized in J. Hawkins (1983: 24–5). Conversely, in the same sample, the languages which show VO and AuxV order are Finnish, Greek, Italian, Mayan, Norwegian, Serbian, and Swahili. This list excludes the VSO languages Welsh, Zapotec, and Masai, all of which have AuxV; if VSO in all these languages is derived as described for Welsh in §1.3.1, then these languages also fit the pattern. Only one language, Guarani, with VAux and VO, diverges from the pattern, although a number of the languages Greenberg studied lacked a class of auxiliaries according to Greenberg's definition of what an auxiliary is (see note 29) and therefore their status regarding this correlation could not be determined.

Correlations of this kind were first observed by Greenberg (1963), and have formed essential data for the field of language typology ever since. They are usually known as implicational universals, since they are often stated in the form of logical implications, i.e. 'if a language has property P, then it has property Q'; we saw an example of an implicational universal in our formulation of Bach's generalization in (113) in the previous section.

What such statements effectively assert is that, of the four combinations of presence vs. absence of properties P and Q which are in principle available, one is not found: presence of P and absence of Q. Here, for example, is the statement of the correlation between OV and VAux (from J. Hawkins (1983: 20)):

In languages with dominant order SOV, an inflected auxiliary always follows the main verb.

If we are to take the idea that parameters of UG are responsible for variation in grammatical systems, then such implicational universals must be derived either from clustering effects created by a single parameter, or by some theoretical articulation of the relations among the values of distinct parameters. In this respect, the results of work in language typology can be seen as setting a particular research agenda for principles-and-parameters theory. For example, let us state parameter F2 as follows:²⁹

F2. Do main verbs precede or follow their auxiliaries in overt order?

PRECEDE: German, Dutch, Japanese, Korean, Turkish, Basque ...

FOLLOW: English, Romance, Zapotec ...

If, in terms of the clause structure we adopted in §1.3, we continue to take auxiliaries to be members of T, then we can conflate F1 and F2 as follows:³⁰

²⁹ Thai is in F1, but not F2. This is because according to Greenberg's (1963: 66) definition of auxiliaries in terms of possible person/number inflection ('a closed class of verbs ... inflected for person and number, ... in construction with an open class of verbs not inflected for both person and number'), Thai cannot have auxiliaries as it lacks such inflections. A different definition of auxiliaries may yield a different result, although the facts in Thai are rather complex. In any case, Thai probably does not constitute a counterexample to the implicational correlation between F1 and F2.

³⁰ The structure we gave in (62) for German does not correspond to F3, as we indicated T as preceding its VP complement. However, we could assume that VP precedes T in German without having to change any fundamental aspect of what was stated there. In fact, as pointed out by Vikner (1995), if T is final in German, then we may assume that V always moves to T, as the finite verb is always final in non-V2 clauses in that language. See also note 15.

Dryer (1992: 98–9) points out that there is no correlation between the relative order of tense/aspect particles and V and that of verb and object; however, as he makes clear in a subsequent section (100–1), this applies to all cases of tense/aspect markers, including affixes, particles, and auxiliaries. If purely verbal auxiliaries are taken into consideration, then the correlation discussed in the text holds for his much larger sample.

- F3. Does the structural complement of V/T precede or follow V/T?
 PRECEDE: German, Dutch, Japanese, Korean, Turkish, Basque ...
 FOLLOW: English, Romance, Thai, Zapotec ...

F3 represents a first step in the direction of **cross-categorical harmony** as a way of understanding and unifying some of Greenberg's word-order universals. This idea originates in Vennemann (1974) and was developed using a variant of the theory of phrase structure assumed here in J. Hawkins (1983.)

A further property which in many languages correlates with F3 is the relative order of adpositions (i.e. pre- or postpositions) and their complements. In Greenberg's thirty-language sample, all but three of nineteen VO languages are prepositional, and all eleven OV languages are postpositional. Dryer (2005b: 386) gives the following figures concerning this correlation in 1,033 languages he sampled (141 languages were defined as not falling into one of the four types):

(125) OV and Po(stpositions)	427
OV and Pr(epositions)	10
VO and Po	38
VO and Pr	417

More than 80 per cent (844) of the languages sampled show consistent orders in this respect, while only forty-eight (less than 4 per cent) of the languages diverge. This is clearly a significant result, although the forty-eight divergent languages require further investigation.

We could thus formulate parameter F4, and unify it with F3 as F5:

- F4. Do objects of adpositions precede or follow their adpositions in overt order?
 PRECEDE: Japanese, Korean, Turkish, Basque, Amharic ...
 FOLLOW: English, Romance, Thai, Zapotec, **German, Dutch** ...
- F5. Does the structural complement of V/T/P precede or follow V/T/P in overt order?

F5 takes us temptingly close to a fully category-neutral statement of word-order variation, which might look like F6:

- F6. For all heads H, does the structural complement of a head H precede or follow H in overt order?

This is fairly close to Dryer's (1992: 116) Branching Direction Theory (BDT), which he states as follows:³¹

³¹ F6 may seem closer to what Dryer (106) calls the Head-Complement Theory, i.e. the theory that verb patterners are complement-takers while object-patterners

Verb patterners are heads and object patterners are fully recursive phrasal dependents, i.e., a pair of elements X and Y will employ the order XY significantly more often among VO languages than among OV languages if and only if Y is a phrasal dependent of X.

Here ‘verb patterners’ and ‘object patterners’ refer to elements which pattern with, respectively, the verb and the object in their relative ordering. Dryer’s BDT is based on a carefully selected sample of 625 languages from all over the world, and as such has impressive empirical scope. (Haspelmath *et al.* (2005) survey a total of 2,560 languages, but not for all the properties given below.)

In terms of what we have seen so far, F5 and F6 present an obvious difficulty concerning German, Dutch, and Latin: these languages apparently pattern one way with regard to F3 and the other way with regard to F4, thereby making the conflation of F3 and F4 given in F5 impossible. According to Dryer (2005b: 388) there are ‘repeated instances’ of the co-occurrence of OV and Prepositions in Iranian languages, including Persian, Tajik, and Kurdish; VO co-occurs with Postpositions in West African languages, some Finno-Ugric languages, and in South America. He adds that languages of both types are ‘often not typical OV or VO languages’ (387).

The expected cross-categorical harmony is not found in cases like German, Dutch, and Latin, then. To deal with this difficulty, we have several options. The first and worst option is to retract any attempt at cross-categorical generalization and revert to the position that the relative order of complement and head is to be restated for each category of head. This is a very weak theory, which would effectively make F3 – and similar correlations to be discussed below – appear to be accidental. It also goes against Dryer’s conclusions, as enshrined in his BDT. An intermediate approach might group heads into subclasses for the purposes of the statement of cross-categorical word-order correlations. We could, for example, consider F3 to be the subclass of F6 which refers to clausal heads, those which make up the ‘core’ sentence TP. Again, Dryer’s results suggest that a stronger theory is called for, although we will adopt this

are complements. His objections to this theory have to do with the relative ordering of verb and manner adverbs, with noun-genitive order and with relatives. Clearly, none of these cases involves complementation if this concept is related to subcategorization, argument structure, assignment of thematic roles, etc. (See §2.3, for more on these notions.) However, F6 refers to structural complementation, i.e. the structure [_{HP} H XP], where H is non-recursive and XP is (potentially) recursive. In this sense, F6 is equivalent to Dryer’s generalization.

option in the next subsection and reconsider it in more detail in §2.5. A third option, pursued in detail in J. Hawkins (1983), is to consider more complex logical relations among the different word-order options. But again, Dryer has shown that many correlations hold among pairs of dyads (for example, OV and Po, as above). Finally, we could look again at the German, Dutch, and Latin data; perhaps the idea that the OV word order seen in subordinate clauses is basic is mistaken. If these languages are taken to be VO and AuxV, they fit into all the generalizations in F1–F6. In different ways this is assumed for German and Dutch by J. Hawkins (1983) and Zwart (1997), and we will see what J. Hawkins says about Latin in the next section. I will reconsider this possibility in §2.5.

Let us finally consider the whole range of word-order correlations established by Dryer. These are as follows (see his Table 39, 108):³²

(126)	a. V	PP	<i>slept [on the floor]</i>
	b. <i>want</i>	infinitive	<i>wants [to see Mary]</i>
	c. Copular	Predicate	<i>is [a teacher]</i>
	d. Aux	VP	<i>has [eaten dinner]</i>
	e. Neg-Aux	VP	<i>don't [know French]</i>
	f. Comp	Sentence	<i>that [John is sick]</i>
	g. Q-marker	Sentence	<i>if [John is sick]</i>
	h. Adverb	Sentence	<i>because [John is sick]</i>
	i. V	Manner Adverb	<i>ran [slowly]</i>
	j. Article	Noun	<i>the [man]</i>
	k. Plural marker	Noun	<i>PL [man] (= 'men')</i>
	l. Noun	Relative Clause	<i>movies [that we saw]</i>
	m. Noun	Genitive	<i>father [of John]</i>
	n. Adjective	Standard of comparison	<i>taller [than John]</i>

All of the correlations in (126) fall under F6, given the assumptions we have made about clause structure, and other fairly plausible assumptions regarding the structure of nominals and the structure of complex adjectival

³² Dryer also shows that the relative order of verb and subject is predicted by the relative order of verb and object. He comments, 'the proportion of the genera [representative language groups – IGR] containing SV languages is higher among OV languages than it is among VO languages, largely because of the extreme rarity of OVS languages' (105). For the reason given, it may not be correct to subsume the correlation under the same principle as the others, i.e. under F6 in the terms adopted here. (See Dryer (1992: 125) for relevant discussion.) I will leave this correlation aside; it is not clear how to capture it in terms of the assumptions about phrase structure and grammatical functions being made here.

constructions. (126a, b) both involve the relative order of verbs and their complements, PP in (126a) and one of CP, TP, or VP in (126b). (126c–e) all plausibly involve the order of T and its structural complement, VP in (126d, e) and either a special PredicateP (see Bowers (1993; 2001)) or a variety of different phrasal complements in (126c). (126f, g) clearly involve the relative order of C and TP, and (126h) may too; alternatively, (126h) involves P with a CP complement (this depends on the exact status of adverbial subordinators as P or C). (126i) is slightly more surprising; however, manner adverbials may be taken to occupy complement positions inside VP, as long as these positions are dissociated from the requirements lexically imposed by heads (*pace* the theory proposed in Chomsky (1981, Chapter 2), for example).

Example (126j) can be understood in terms of the DP-hypothesis (which we briefly mentioned in the Introduction), according to which articles are members of the category D, which has an NP complement. Similarly, (126k) may reflect an elaboration of that hypothesis which claims that D takes NumberP as its complement, and NP is the structural complement of Number. (We will see this idea again in a different context in §2.2.) (126l) is consistent with the idea that the relative clause (a CP) is the structural complement either of the head noun, or of the D which introduces the entire structure (see Kayne (1994); Bianchi (1999) on the latter idea). (126m) is straightforward for some instances of genitives (for example, the one given as illustration; here *of John* is uncontroversially the complement of *father*), but not for others (for example, *Caesar* in *Caesar's destruction of the village*). Finally, (126n) can be maintained if the 'standard of comparison' is taken to be the (CP) complement either of the comparative/superlative morpheme or of *than*. (Bhatt and Pancheva (2004) give a recent version of this analysis, which has its origins in Chomsky (1965:178ff).)

All of the above points raise analytical questions, but they are at least compatible with the idea that a parameter like F6 may underlie Dryer's results. English shows a very consistent VO pattern, except that genitives may precede their nouns (as in *John's father*); in French and other Romance languages, where this kind of genitive is not available, the VO pattern is completely consistent with the VO value of F6. We can illustrate a consistent OV pattern with the Dravidian language Malayalam. To the extent that the phenomena listed in (126) are found, Malayalam instantiates a consistent OV pattern, as shown in (127) (all examples from Asher and Kumari (1997), with page references as indicated; capitals indicate retroflex consonants):

- (127) a. **mee** **ʃayuTe** **meele/miite** **vecciTTiTūNTə** (PP V)
 table-Gen on put-Perf₂-Pres
 ‘put on the table’
 (123)
- b. **vi** **ʃvasikka** **vayya** (infinitive V)³³
 believe-Inf₁ cannot
 ‘cannot believe’
- c. **Ii** **kuTTi** **nallavan** **aaNə**. (Predicate Copula)
 this child good-Masc be-Pres
 ‘The child is good.’
 (97)
- d. **Nii** **atə** **ceyyaNam**. (VP Aux)³⁴
 you it do-Deb
 ‘You must/should do it.’
 (153)
- e. **Nii** **atə** **ceyyaNTa**. (VP Neg-Aux)
 you it do-Deb-Neg
 ‘You need not do it.’
 (153)
- f. **Ellaarum etti** **ennə** **kuTTi** **paraññu**. (Sentence C)
 all arrive-Past QP³⁵ child say-Past
 ‘The child said that all had arrived.’
 (47)
- g. **Avan vannoo?** (Sentence Q)
 he come-Past-InterrogativePrt
 ‘Did he come?’
 (8)
- h. **Ñaan parañña** **poole** **avan** **pravartticcu**. (Sentence Adverb)
 I say-Past-RelPpl like he act-Past
 ‘He acted in the way I told him.’
 (83)
- i. **AvaL** **bhamgiyaayi** **prasamgiccu**. (Manner-Adv V)
 she beauty-Adv speak-Past
 ‘She spoke beautifully.’
 (112)

³³ An example with ‘want’ was unavailable, but it is well-known that in many languages the complement of ‘can/be able to’ is syntactically very similar to that of ‘want’, and so this example may suffice.

³⁴ Malayalam being an agglutinating language, it is unclear whether the relevant morphemes in this example or the next are endings or auxiliaries. It is nevertheless clear that they follow the VP, and so they may illustrate the relevant property.

³⁵ ‘Quotative Particle’ rather clearly a kind of complementizer, to judge by its distribution (Asher and Kumari 1997: 45ff.).

- j. There is no separate category of articles. (Noun Article)
(124)
- k. Number is marked by a suffix. (Noun Plural)
(249)
- l. **naaLe** **naTakkunna** **ulsavam** (RelCl Noun)³⁶
tomorrow take-place-Pres-RelPpl festival
'the festival that takes place tomorrow'
(54)
- m. **aa** **kuTTiyuTe** **peena** (Genitive Noun)
that child-Gen pen
'that child's pen'
(132)
- n. Raaman **kṛṣṇane** **kaalum miTukan** **aaNə.** (Standard Adj)
Raman Krishnan-Acc than clever be-Pres
'Raman is cleverer than Krishnan.'
(169)

Thus Malayalam represents the OV value of F6, while English (with the proviso for genitives mentioned) and the Romance languages represent the VO value.

There is, however, a problem with F6. Dryer shows that a minority of the languages in his sample actually conform to the BDT on all points. The majority of languages diverge at least in some respects (the commonest divergence being N-Rel order in OV languages). Thus, if F6 is a single parameter, it predicts a spectacular clustering of properties, which is not actually attested in the majority of languages. In order to solve this problem, we need to be clearer about how the linear orders relevant to F6 are produced. F6 can only be a true parameter if it is associated with a unitary grammatical operation, for example, whatever operation determines which branch of binary-branching pair is the recursive one. (Such a formulation would be the closest parametric analogue to Dryer's BDT.) However, it may be that linear order is the result of the interaction of various grammatical operations. If this were the case, then we may be able to regard a more abstract version of F6 as a 'pure' word-order parameter, and then appeal to other operations which are independently parameterized in order to capture the many divergences from the two 'pure' patterns given by F6. The preference for 'harmonic' ordering may thus derive from an overriding tendency for independent parameters to conspire to produce a certain type of

³⁶ There is another relative-clause construction involving wh-movement, in which the head noun appears to precede the clause – see Asher and Kumari (1997: 53). However, the participial construction illustrated here is more common.

grammar. This implies that there may be a higher-order cross-linguistic principle at work, in addition to parameters as we have described them in this chapter. This approach is most in line with both J. Hawkins' notion of cross-categorial harmony and Dryer's BDT, and we will develop it in §2.5 and §3.5.

In this section I have outlined the idea that systematic cross-linguistic variation in head-complement orders exists, at least across significant subsets of heads. I have attempted no theoretical explanation of this, beyond simply observing that cross-categorial harmony, in essentially the sense introduced by J. Hawkins (1983), and developed by Dryer (1992), is a promising idea for the application of the theory of principles and parameters in this domain. For more details on this and nearly all the other issues raised here, the reader is referred to the works cited, in particular Song (2001); Croft (2003).

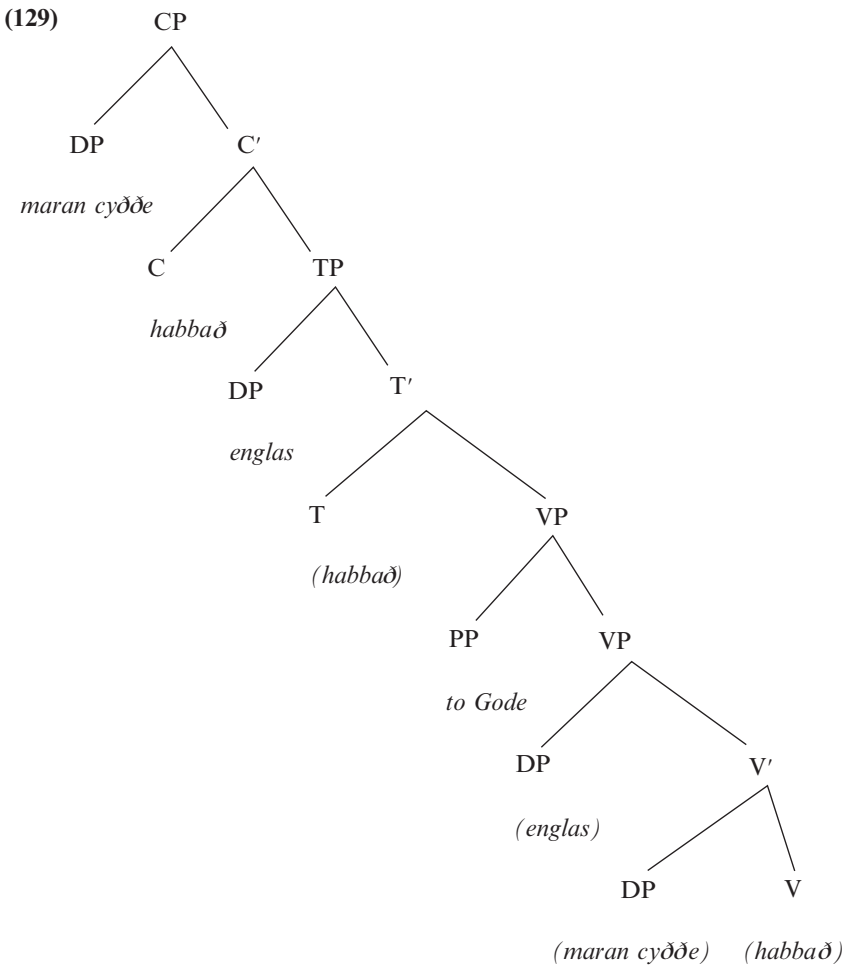
1.6.2. *Head-complement order diachronically*

In this section I will focus simply on OV and VO orders, limiting my attention to the restricted cases F1 and F3 discussed in the previous section (although the drawbacks of these more restricted formulations have been noted). We already made the basic observation in the Introduction: Old English showed OV word order in embedded clauses. We mentioned this in our discussion of verb second in §1.3.2, and we repeat example (78) from that discussion here:

- (78) a. ... þæt ic **þas boc** of Ledenum gereordre to Engliscre spræce **awende**
 ... that I this book from Latin language to English tongue translate
 '... that I translate this book from the Latin language to the English tongue'
 (AHTh, I, pref, 6; van Kemenade 1987: 16)
- b. ... þæt he **his stefne up ahof**
 ... that he his voice up raised
 '... that he raised up his voice'
 (Bede 154.28)
- c. ... forþon of Breotone **nædran** on scippe **lædde wæron**
 ... because from Britain adders on ships brought were
 '... because vipers were brought on ships from Britain'
 (Bede 30.1–2; Pintzuk 1991: 117)

We can thus observe that OE had OV order in subordinate clauses. We saw in §1.3.2 that OE had verb-second order in main clauses, and so we do not expect to find overt OV order in such clauses (unless of course the object is fronted to first position). An object-initial verb-second clause in OE like (128) can thus be treated as having a structure just like that proposed for such clauses in German in (62):³⁷

- (128) Maran cyððe habbað englas to Gode þonne men.
 more affinity have angels to God than men
 ‘Angels have more affinity to God than people.’
 (AHTH, I, 10; van Kemenade 1987: 42)



³⁷ Here I have placed the indirect object *to Gode* in a position adjoined to VP. This is an approximation of the standard analysis of indirect objects (on which see

The position of the auxiliary in (79c) indicates a further parallel with German. Here we see the finite auxiliary in a subordinate clause following the non-finite verb. (See note 30 on one possibility for analysing this order in German, which would carry over to OE.) So we conclude that parameter F3 had the value PRECEDE in OE, and that it has therefore changed in the subsequent history of English.

While the situation concerning F3 is reasonably clear, that regarding F6 is not. J. Hawkins (1983: 335) states that OE is prepositional, with Article-Noun, Genitive-Noun and Noun-Relative order. To this we can add that the complementizer *þæt* always precedes the sentence it introduces. So OE is clearly mixed as far as F6 is concerned, and more needs to be said.

OE word order was rather freer than that of Modern German, an observation which has given rise to much discussion. In particular, non-pronominal direct objects could follow the non-finite verb as well as precede it:

- (130) a. *þæt hi urum godum geoffrian magon ðancwurðe onsægednysse*
 that they our gods offer may grateful sacrifice
 ‘that they may offer a grateful sacrifice to our gods’
 (*ÆCHom* I, 38.592.31; Fischer *et al.* 2000: 144)
- b. *ðe is genemned on Læden Pastoralis, & on Englisc Hierdeboc*
 which is named in Latin Pastoralis and in English Shepherd’s-Book
 ‘which in Latin is called Pastoralis and in English Shepherds’ Book’
 (*CPLetWærf* 58; Fischer *et al.* 2000: 144)

In (130a), the auxiliary follows the main verb, while in (130b) the auxiliary precedes the main verb. We also find examples where the auxiliary precedes the non-finite verb (in contexts which are clearly not those of the general verb-second rule – see §1.3.1) but where the object precedes both verbal elements, as in:

- (131) a. *se ðe nan ðing nele on ðissum life ðrowian*
 he who no thing not-wants in this life suffer
 ‘he who will suffer nothing in this life’
 (*ÆCHom* I, 10.164.22; Fischer *et al.* 2000: 52)
- b. *Gif he ðonne ðæt wif wille forsacan*
 if he then the woman wish refuse

in particular Radford (1997, Chapter 9); Radford (2004: 345); Hornstein *et al.* (2005: 92ff.) and §2.3, and should not be interpreted as indicating a formal parallelism with the adverbial phrase *schon letztes Jahr* in (62). I have omitted *þonne men* from the representation as it obscures the parallel with (62), and it is rather unclear where this constituent (which is most likely a CP containing a great deal of elided material) should attach.

‘if he then wishes to refuse the woman’
(*CP* 5.43.15; Fischer *et al.* 2000: 52)

This variation in word order presents a further challenge to the postulation of a parameter like F3, which we will come back to in §2.5. Nevertheless, it is fairly clear that the OV pattern is the statistically predominant one in OE. See Pintzuk (2002) for discussion of this and a number of related points, in particular regarding the relation between surface OV or VO order and other variant properties, such as the position of particles, etc.; Pintzuk interprets the variation in OE word order as evidence that there is more than one grammar underlying the corpus.

Dating the change in parameter F3 in English is rather difficult. On the one hand, there is evidence for VO and AuxV order in OE, as we have already seen. On the other hand, superficial OV order can be found in the fifteenth century and even, in the relevant context and the right register, as late as the seventeenth century. Once again, this points perhaps to a non-unitary parameter. I reserve judgement on this question here; we can simply observe that the change took place during the ME period. The general question of the gradualness or otherwise of parameter changes will be discussed in more detail in §4.1.

The same kind of word-order change can be observed in the development from Latin to Romance. Classical Latin word order, although rather free, is tendentially OV. See Vincent (1988: 59ff.); Harris (1978: 18ff.) and the references given there; Ernout and Thomas (1993: 161). The following example illustrates:

- (132) a. Ego ... **apros** tres et quidem pulcherrimos cepi.
I boars three and indeed very-beautiful have-taken
‘I have taken three indeed very beautiful boars.’
(Pliny the Younger)
- b. Caesar Aeduos **frumentum flagitabat**.
Caesar Aedui corn was-demanding
‘Caesar kept demanding the corn of the Aedui.’
(Vincent 1988: 59)

In Latin, it was also usual for auxiliaries to follow main verbs. This is true to the extent that a class of auxiliaries can be discerned; the clearest case involves *esse* (‘be’) in the perfect passive or perfect of deponent verbs, as in (134):

- (134) a. ut ... ad ciuitatem gemitus popoli omnis **auditus sit**
that to city groan of-people all heard be

‘that the groans of all the people be heard (as far as) the town’
(*Peregr. Aeth.* 36, 3; Ernout and Thomas 1993: 229)

- b. *Cyrene autem condita fuit ab Aristaeo.*
Cyrene however founded was by A.
‘Cyrene was however founded by Aristaeus.’
(Justin 13, 7, 1; Ernout and Thomas 1993: 229)

So here we see evidence for the PRECEDE value of parameter F3, as in OE. And again, it is clear that the Modern Romance languages have the FOLLOW value, just like Modern English, as the following French translations of (132a) and (133b) show:

- (134) a. *J’ai pris trois très beaux sangliers.*
I’ve taken three very beautiful boars.
b. *Puis Cyrène a été fondé par A.*
Then Cyrene has been founded by A.

It is not entirely clear when this change took place in the history of Latin/Romance, although Harris (1978: 19) suggests that it may have happened in the Vulgar Latin period, by the fourth or fifth century AD.

As with OE, the other features of Latin are mainly VO (see Harris (1978: 19ff.)). J. Hawkins (1983: 331) states that it is prepositional and has Noun-Relative order, with either order of Noun and Genitive. Moreover, it is clear that complementizers and question words such as *ut* and *ne* precede their clauses:

- (135) a. *ad Romam* (Pr)
to Rome
b. *Germani qui trans Rhenum incolunt* (N-Rel)³⁸
Germans who across Rhine lived
‘Germans who lived across the Rhine’
(Caesar, *Bello Gallico* I, 1, 3; Ernout and Thomas 1993: 335)
c. *liber Petri* (N-Gen)
book of-Peter
‘Peter’s book’
d. *Ubii Caesarem orant ut sibi parcat.* (C – Sentence)
Ubii Caesar beg that them(elves) he-spare

³⁸ Ernout and Thomas (1993: 333) also give the following example, which appears to have N-Rel order:

- (i) *quas Numestio dedi litteras*
which to-Numestio I-gave letters
‘the letters which I gave to Numestius’
(Cicero, *At.* 2, 24, I)

‘The Ubii beg Caesar to spare them.’
(Vincent 1988: 66)

So we again observe a mixed picture as regards F6, but a clear change in parameter F3.

We can infer or observe a similar OV-to-VO change elsewhere in Indo-European. Rögnvaldsson (1996) and Hróarsdóttir (1996; 1999; 2000) show in detail that Old Icelandic was OV. The relevant examples for parameter E3 are:³⁹

- (136) a. so þorsteinn skyldi lífinu tapa
so Thorstein should life-the lose
‘so that Thorsteinn should die’
b. þú munt fret hafa, að ...
you will heard have, that ...
‘you will have heard that ...’ (Hróarsdóttir 1999: 298)

Indeed, OV order is usually assumed for early Germanic generally (see the discussion in J. Hawkins (1983: 221ff.) and the references given there). A. Taylor (1994) argues that Greek changed from OV in the Homeric period to VO in the Classical period. Early Old Irish shows some OV orders. (See Bergin (1934–8); Thurneysen (1946: 327); Russell (1995: 286ff.); and the more recent discussion in Doherty (2000a, b)); Continental Celtic also shows some evidence of OV order (Russell 1995: 282ff.) On the other hand, the modern Celtic languages are all VO, in fact VSO. It seems, then, that changes in parameter F2, and possibly F3, are widespread, at least in the history of Indo-European. Indeed, W. Lehmann (1993: 190) cites Delbrück (1893–1900) for proposing ‘OV order for the early dialects of Indo-European and the parent language as well’. Fortson (2004: 142) says that ‘[i]t is almost universally asserted that most of the ancient IE languages were verb-final, and that PIE [Proto-Indo-European – IGR] was as well’ (although he goes on to point out that this claim ‘needs tighter formulation and convincing demonstration’). If this is correct, then all present-day Indo-European languages with VO order (English, North Germanic, Romance, Celtic, Greek, Slavonic, etc.) must have undergone a change in parameter F2, at least. W. Lehmann (1993: 203–5) also cites evidence that Indo-European may have been postpositional

³⁹ Both Rögnvaldsson (1996) and Hróarsdóttir (1996; 1999; 2000) use the term ‘Old Icelandic’ for the language which is frequently referred to as ‘Old Norse’. See Rögnvaldsson (1996: 56) and Faarlund (1994b: 38) for a clarification of the terms for the older stages of the Scandinavian languages.

and had Standard-Adjective and Relative-Noun orders. This suggests a more far-reaching OV typology than can be ascertained for OE or Latin. A great deal has been said about Indo-European syntax, and we will return to this topic in §4.4, when we discuss syntactic reconstruction. The OV-to-VO change is not, however, restricted to Indo-European: Kiparsky (1996: 172) observes that ‘Western Finno-Ugric languages, including Finnish’ have undergone a change from OV to VO.

The above observations show that, despite certain difficulties in interpreting the data and despite the problematic nature of parameter F6, there is good evidence that parameter F3 has changed in English, Icelandic, and Romance, and reason to think that these parameters may have also changed in other branches of Indo-European and elsewhere.

1.7. Summary

Here I will simply summarize the main conclusions regarding the diachronic changes in parameter values that we have seen in this chapter. The goal of this chapter has been to demonstrate that the notion of parameter, as it has been construed in work on comparative syntax since Chomsky (1981), can play a useful role in describing syntactic variation in the diachronic dimension, just as it can in the synchronic domain. In other words, parameter change may be, at the very least, a useful analytical tool for diachronic syntax. Here are the changes we have seen:

A: The null-subject parameter: changed value in French ca. seventeenth century, also in Northern Italian dialects at about the same time, and presumably in prehistoric Germanic; it may be changing in some varieties of Brazilian Portuguese. (See Duarte (1993; 1995), Barbosa, Duarte, and Kato (2005); the papers in Kato and Negrão (2000); §4.2.6, and the references given there), as well as in Welsh (Tallerman (1987).)

B: V-to-T parameter: changed in English ca. 1600, also in Danish and Swedish ca. 1400; appears to have changed in many French-based Creoles (see §5.3, and the references given there)

C: V2 parameter: changed in English in the fifteenth century, in French ca. 1600, also in most, if not all, other Romance languages at various points in the Medieval period; in Welsh ca. 1600; possibly in pre-Old Irish (Doherty 2000a, b).

D: Negative-concord parameter: changed value in French ca. 1600. Other Romance languages (for example, Italian, Spanish, and Portuguese, but not Catalan) may have developed similarly to French.

E: Wh-movement parameter: changed between the Nara and Heian Periods of Old Japanese. *Wh-in-situ* appears in main clauses in recent colloquial French and in Brazilian Portuguese.

F: Head-complement parameter: case F3 of this parameter changed in ME, at some stage in the recorded history of Icelandic, and in Vulgar Latin, also possibly between Homeric and Classical Greek (800–500 BC), and presumably in Celtic, as well as elsewhere in Indo-European, if W. Lehmann (1993) is correct.

On the basis of what we have seen in this chapter, then, there can be no doubt that the notion of parametric change has a role to play in the study of diachronic syntax, and a central role at that. However, this conclusion raises a number of questions which I have so far been avoiding. Foremost among these is the one summarized in the following quotation:

If we isolate a parametric difference between, say, English and Italian, then we can simply describe the parameter and its consequences, and ideally say something about the typology it implies and the trigger experience that sets it . . . Then our job is done and we can go to the beach . . . If, however, we isolate a parametric difference between one historical stage of English and another, then we need to explain not just what the parameter is and what its effects are, but *how, at some point in the generation-to-generation transmission of language, the new value was favoured over the older one.*

(Roberts 1996: 280, emphasis in original).

The question of how parameters can change their value in the generation-to-generation transmission of language is a very difficult and intriguing one. It is also of some importance for the theory of principles and parameters, since a successful answer to this question is likely to tell us much about the nature of parameters, the kind of primary linguistic data required to set them to a particular value, whether there are default values, and potentially many other matters. Much of what follows is devoted to fleshing out and considering these consequences of the point that changes in the values of parameters are an important force in the historical syntax of languages.

Further reading

The history of French

Thurneysen (1892) is the original study in which the observation that null subjects could only appear in V2 clauses was made. **Vanelli, Renzi, and Benincà (1985)** was the first paper to clearly show that Old French and the Medieval Northern Italian dialects had a V2 system with non-clitic subject pronouns and null subjects, which later developed into a non-V2 system with clitic subject pronouns and no null subjects. **Adams (1987a, b)** were the first works in government-binding theory to analyse the OF pattern of combined V2 and null subjects just in V2 contexts, and they were influential for much subsequent work. **Adams (1988a, b, c)** dealt with some of the problems left unresolved in the earlier work, particularly concerning early OF. **Vance (1988; 1997)** looks in detail at Middle French, arguing that the changes affecting V2 and null subjects may not be a single change. **Roberts (1993a)** extends Adams' analyses and attempts to unite a number of changes in the history of French as a single parameter change. There is also a chapter on the development of auxiliaries and the loss of V-to-T movement in English. Further studies of these and related phenomena in the history of French are **Dupuis (1988; 1989)**; **Hirschbuhler (1990)**; and **Hirschbuhler and Junker (1988)**. **Ayres-Bennett (1996)** is a very useful collection of texts from various periods in the history of French, with commentary on all structural features, including morphology and syntax. **Ayres-Bennett (2004)** is primarily a study of attitudes to language in seventeenth-century France; since French was undergoing a number of major changes at that time, much careful documentation of the textual evidence for the grammatical changes is adduced. **Harris (1978)** is a wide-ranging study of the development of French, as well as to a lesser extent Spanish and Italian, from Latin, using a loosely Greenbergian descriptive-typological framework. **De Kok (1985)** is a detailed study of the distribution of complement clitics in Old French. **Foulet (1921)** is a detailed descriptive study of the development of the syntax of interrogatives in the history of French. He notes that *wh-in-situ* first appears in the nineteenth century. **Price (1971)** is a descriptive historical treatment of the development of Latin into French. **Einhorn (1974)** is a general history of French. **Foulet (1990)** is the standard reference for the descriptive syntax of Old French. **Robert (1992)** is a comprehensive historical dictionary of French.

Null subjects

Rizzi (1982) contains a number of classic articles in principles-and-parameters theory. In particular, in Chapter 4 he presents a highly influential account of the null-subject parameter, in which he argues for the first time that wh-movement of a subordinate-clause subject across a complementizer moves the wh-phrase from postverbal position. **Rizzi (1986a)** presents the standard government-binding analysis of null subjects and (arbitrary) null objects. **Alexiadou and Anagnostopoulou (1998)** is an influential proposal concerning the analysis of expletive null subjects in the context of the version of minimalism in **Chomsky (1995, Chapter 4)**, which much subsequent work proposes carrying over to referential null subjects. The latter idea is anticipated and developed at length in **Barbosa (1995)**. The same basic idea was first developed in government-binding terms by **Borer (1986)**. **Barbosa, Duarte, and Kato (2005)** is a systematic study of a number of apparent differences in the distribution and interpretation of null subjects between European and Brazilian Portuguese, partly based on the thorough diachronic studies of Brazilian Portuguese in **Duarte (1993; 1995)**. (We will return to these issues in §4.2.6.) The papers in **Kato and Negrão (2000)** also concentrate on Brazilian Portuguese. **Haegeman (2000)** is a detailed study of the phenomenon of ‘diary drop’ – apparent null subjects in a stylistically restricted register – in English and French. **Cardinaletti (1990)** is an early analysis of expletive null subjects in German. **Roussou and Tsimpli (2006)** present a novel analysis of VSO orders in Modern Greek.

Subject clitics

Brandi and Cordin (1989) is an early study of subject clitics in Trentino and Fiorentino, in which they establish that the subject clitics in these dialects are not pronouns but agreement markers, and that wh-movement of a subject over a complementizer moves an inverted subject only. **Rizzi (1986b)** presents further arguments that subject clitics in many Northern Italian dialects are agreement markers rather than pronouns. **Poletto (1995)** provides an account of the development of subject pronouns into subject clitics in the history of Veneto. **Poletto (2000)** is an analysis of the nature and positions of subject clitics in over a hundred Northern Italian

dialects. **Cardinaletti and Repetti (2003)** argue that the subject clitics of at least some Northern Italian dialects are pronouns, rather than agreement markers. **Jaeggli (1982)** is the earliest government-binding analysis of clitics and null subjects in Romance, concentrating on Spanish and French. He argues that the French subject pronouns are agreement markers. **Roberge (1990)** also argues that in at least some varieties of French the subject pronoun is in fact an agreement marker, as does **Sportiche (1998)**. **Renzi (1983)** looks at subject clitics in eighteenth-century Florentine. **Cardinaletti and Starke (1999)** present a detailed typology of pronouns, dividing them into strong, weak, and clitics. They motivate this typology in terms of a theory of structural deficiency. **Roberge and Vinet (1989)** is a collection dealing with clitics and related matters in a range of Romance varieties.

More general works on Romance syntax

Kayne (1975) is the ground-breaking work in comparative generative syntax, being the first in-depth study of a language other than English. This work set the agenda for work on French and the Romance languages more generally for many years to come, and led to the development of the principles-and-parameters approach. **Kayne (1984)** is a collection of articles from the period 1979–84 mainly on the comparative syntax of English and French; a number of important and influential innovations in early government-binding theory are proposed here. **Kayne (1989)** is a detailed study of clitic-climbing, proposing a specific connection between this phenomenon and null subjects. **Kayne (1991)** relates infinitive movement and enclisis in Romance to the null-subject parameter and to the possibility of certain types of infinitive structure. **Kayne (1994)** proposes the very important antisymmetric theory of phrase structure, which, as we shall see in detail in §2.5.4, has the consequence that all languages have the same underlying word order. **Kayne and Pollock (1978; 2001)** both propose analyses of the French ‘Stylistic Inversion’ construction, assuming rather different theoretical backgrounds in the two papers. **Pollock (1986)** is a further study of Stylistic Inversion in which he insightfully connects the phenomenon to expletive null subjects.

General aspects of generative theory

Chomsky (1957) is the earliest monograph on generative grammar, and remains in many ways a classic exposition of the theory. The analysis of English auxiliaries presented there remains influential. **Chomsky (1959)** is a review of Skinner's *Verbal Behavior*, a behaviourist account of first-language acquisition. The review effectively destroyed behaviourist accounts of language acquisition, at least in linguistics, and went a long way towards establishing a nativist alternative. **Chomsky (1981)** is the foundational text of government-binding theory, and gives the first clear overview of how the modules of that theory (binding theory, Case theory, bounding theory, etc.) were thought to interact. **Chomsky (1982)** introduces some refinements, notably the Extended Projection Principle. **Chomsky (2001)** is, at the time of writing, the fullest exposition of the current version of minimalism. This is what the technical notions introduced in this book are loosely based on. **Chomsky (2004)** further refines and develops this model, while **Chomsky (2005a, b)** discusses the conceptual background to and implications of the current theory. **Hauser, Chomsky, and Fitch (2002)** is an important recent paper in which it is suggested that certain aspects of the human language faculty have correlates in cognitive abilities in other animals. Aspects of 'narrow syntax' (the formal operations of syntax, especially Merge) may, however, be uniquely human. **Pinker and Jackendoff (2005)** and **Jackendoff and Pinker (2005)** are responses, arguing that the earlier view of a dedicated language module of cognition is correct. **Smith and Tsimpli (1995)** is a study of a linguistic savant, an individual in whom linguistic abilities appear to be hypertrophic. They argue that this supports the view that there is a dedicated cognitive module for language. **Anderson and Lightfoot (2002)** is a general introduction to Chomskyan linguistics, which includes a useful chapter on language change. **Pullum and Scholz (2002)** present a very strong version of the argument from the poverty of the stimulus, which they then argue must be false. **Jackendoff (2002)** is a general introduction to generative linguistic theory with emphasis on the place of linguistic theory in cognitive science.

Verb-movement

Den Besten (1983), written and first circulated in 1977, remains the classic treatment of verb-movement to C, unifying the analysis of Germanic V2,

French subject-clitic inversion and English subject-aux inversion. **Emonds (1978)** is an early study of verb-movement, in which what was later analysed (in the terminology used here) as V-to-T movement was first identified. **Pollock (1989)** is the fundamental article on verb-movement, systematically showing how English and French differ in this respect. This article also introduced the ‘split-INFL’ clause structure, which ultimately led to more elaborate structures such as that proposed by Cinque (1999). **Belletti (1990)** extended Pollock’s (1989) analysis of verb placement to Italian, showing that Italian infinitives move to a higher position than their French counterparts. **Travis (1984)** is an important and influential study of V2 and expletive constructions in the Germanic languages, noted for the fact that she argued, *contra* den Besten (1983), that subject-initial V2 clauses are TPs. **Vikner (1995)** is a comprehensive study of verb-movement and transitive expletive constructions across the Germanic languages. **Schwartz and Vikner (1996)** argue that the verb always leaves TP in V2 clauses, *contra* Travis (1984) and Zwart (1997). **Zwart (1997)** presents an early minimalist analysis of verb-movement in Dutch, in which he argues, in agreement with Travis (1984), that subject-initial V2 clauses are TPs. **Roberts (1999)** looks at the status of V-to-T in various French-based creoles; we will look at this in more detail in §5.3.2. **Cinque (1999)** is an important and influential study of adverbs, functional categories and clause structure, in which he advocates that simple clauses may contain more than forty functional categories. **Cinque (2004)** applies these results to the analysis of restructuring and clitic-climbing phenomena in Romance. **Emonds (1980)** features the first version of the analysis of VSO presented in §1.3.1. **Carnie and Guilfoyle (2000)** is a collection of papers on VSO and VOS languages. **Massam (2000)** proposes an analysis of VOS and VSO orders in the Polynesian language Niuean in terms of (remnant) VP-fronting. **Rackowski and Travis (2000)** is an analysis of Malagasy and Balinese, taking into account also Niuean, of the same type. **Shafer (1994)** is a comprehensive discussion of word order and clause structure in Breton. **Jouitteau (2005)** is a more recent study of word order, agreement, and related phenomena in Breton. **Willis (1998)** is a detailed study of the loss of V2 in Welsh. **Shlonsky (1997)** similarly studies word order and clause structure in Hebrew and a range of Arabic dialects. **Thurneysen (1946)** is the most comprehensive grammar of Old Irish to date, and the main source of information on that language. **Bergin (1938)** put forward what has become known as ‘Bergin’s Law’, an important observation concerning Old Irish syntax which states that when the verb is not in

initial position in the clause, it must take on a particular form of inflection, the ‘conjunct’ inflection. **Russell (1995)** is a comprehensive introduction to Celtic philology, with an interesting chapter on the development of VSO and the complex double-inflectional system of Old Irish. **Doherty (2000a, b)** analyses Bergin’s Law and related phenomena concerning the position and inflection of the Old Irish verb from a minimalist perspective.

Negation

Ladusaw (1980) is the first systematic study of negative-polarity items in modern formal semantics. He proposes an important generalization concerning the semantic nature of the context in which these items appear, a generalization challenged in Giannakidou (1997; 1998). **Giannakidou (1997; 1998; 2000)** are in-depth studies of negation, negative concord and negative-polarity, primarily from a semantic point of view, although there is some syntactic analysis too. **Horn (2000)** summarizes the issues arising in connection with the analysis of both free-choice and negative-polarity *any*. **Horn and Kato (2000)** is a collection of articles on the general topic of the syntax and semantics of negation. **Watanabe (2004)** gives an analysis of negative concord in Japanese and other languages in terms of Agree, and draws different conclusions on the nature of n-words from Giannakidou (2000). **Zeijlstra (2004)** also analyses negative concord in terms of Agree, applying the analysis to a wide range of languages, as well as looking at Jespersen’s Cycle and other diachronic changes affecting negation. **Déprez (1997; 1999; 2000)** are detailed treatments of French and Haitian Creole negation. **Martins (2000)** analyses negative concord and n-words synchronically and diachronically across a range of Romance languages. **Borsley and Morris-Jones (2005)** is an analysis of Welsh negation. **Jespersen (1917)** famously proposed the cycle of changes in clausal-negation marking, which has since become known as Jespersen’s Cycle.

Wh-movement

Ross (1967) is the classic study in which island phenomena were first presented and analysed. **Chomsky (1973)** is the foundational article of the Extended Standard Theory of generative grammar, the version of the

theory current in the 1970s. It contains the first statement of the subadjacency condition. **Huang (1982)** is the classic work on wh-expressions in Chinese, showing that, while they are *in-situ* elements, facts concerning scope and selection strongly suggest a ‘covert movement’ analysis. **Watanabe (2001)** provides an overview of research on wh-*in-situ* languages. **Bach (1971)** is one of the earliest discussions of wh-*in-situ*; this is where Bach’s generalization was first put forward. **Cheng (1991)** is an important study of cross-linguistic variation in wh-constructions, in which the importance of clause-typing is brought to the fore. **Rudin (1988)** was the first analysis of multiple wh-movement in Slavonic. **Bošković (2002)** puts forward an analysis of the difference between single and multiple wh-movement languages which does not treat the latter as a parameter dependent on the positive setting of parameter E. Instead, he suggests that parameter E also distinguishes among multiple-fronting languages, and that what distinguishes the latter from single-movement and *in-situ* languages is a further requirement that all wh-phrases must be focused. Multiple-movement thus involves a single instance of wh-movement and multiple-focus-movement. **McCloskey (2001)** puts forward an analysis of the alternating complementizers in Irish, apparently triggered by wh-movement from the clause introduced by the complementizer. **Watanabe (2002)** provides the account of the development of wh-*in-situ* in Old Japanese that was summarized in §1.5.2, while **Watanabe (1996)** provides an analysis of wh-agreement in Palauan and other languages. **Mathieu and Sitaridou (2005)** present an analysis of the loss of wh-movement from a left-branch in the history of Greek. **Rossi (1993)** looks at the emergence of the possibility of wh-*in-situ* in Brazilian Portuguese. **Rizzi (1990; 2000)** present a general theory of locality of wh-movement and other operations, known as relativised minimality.

Language acquisition and parameter-setting

Clark and Roberts (1993) attempts to develop a general account of how parameter values are set and changed on the basis of primary linguistic data; the account is applied to the changes in the history of French discussed in §1.1.2. **Dresher (1999)** presents a cue-based learning theory for phonological parameters, and shows how a natural ‘learning path’ emerges. We will return to these ideas in detail in Chapter 3. **Guasti (2002)** is a recent textbook

on language acquisition, which gives a thorough overview of work to date on first-language acquisition from the principles-and-parameters perspective. We will refer to this work extensively in Chapter 3.

Language typology

Greenberg (1963) is the foundational text in language typology, where, on the basis of a thirty-language sample, forty-five syntactic and morphological universals of various types were tentatively proposed. Some of these have not been falsified by much larger recent surveys. **W. Lehmann (1973)** was the first attempt to apply Greenbergian typology to word-order change. This article was also the first to propose that OV and VO are the most important predictors of other aspects of word order. **Vennemann (1974)** developed Lehmann's ideas, proposing the Natural Serialization Principle as an account of Greenbergian word-order correlations. **J. Hawkins (1983)** was a major contribution to language typology, where, among other things, the notion of cross-categorical harmony was developed in terms of X'-theory. **Dryer (1992)** is a major study, the most comprehensive of its time, of the Greenbergian word-order correlations. On the basis of a sample of 625 languages from all over the world, the Branching Direction Theory described in §1.6.1 is proposed. **Haspelmath et al. (2005)** is the most comprehensive survey of the structural features of the known languages of the world to date, covering 141 phonological, morphological, syntactic and lexical features in over 2,560 languages. The database is also available as a CD. **Haspelmath (1997)** is a thorough and interesting typological study of indefinite pronouns. **Comrie (1989)** is a classic introduction to language typology. **Song (2001)** and **Croft (2003)** are more recent and more detailed introductions.

The history of English

van Kemenade (1987) was the first detailed generative study of Old English syntax. It is clearly shown that many of the central features of OE syntax are similar to or identical with those in Dutch and/or German. **Kroch (1989)** introduces the important concepts of the Constant Rate Effect and grammars in competition, applying both of them to a number of case studies

including notably the development of *do*-support and the loss of V-to-T movement in Early Modern English. **Pintzuk (1991; 1999)** applied Kroch's (1989) idea of grammars in competition to word-order variation in Old English. In **Pintzuk (2002)**, these ideas are developed further; in particular it is argued that quantified objects undergo a special leftward-movement rule. **Kiparsky (1996)** proposed the first account of word-order change which assumed Kayne's (1994) universal underlying word order. **Fuß and Trips (2002)** also propose an account of word-order change in English, which makes use of a restricted version of the competing grammars idea. (We return to the topic of competing grammars in more detail in §4.1 and §4.2.) **Haeberli (2002)** is a study of Old English word order with special emphasis on V2 and the nature of the subject position. **Jonas (1996)** is a study of verb-movement, transitive expletive constructions and related matters in Faroese and in the history of English. **Warner (1997)** provides an overview and attempted synthesis of the accounts of the loss of V-to-T movement in Early Modern English available at the time. He also provides a detailed chronology, which is only partly compatible with the results of Kroch (1989). We will return to this in §2.1. **Kroch and A. Taylor (1997)** propose an account of the loss of V2 in Middle English which relies on the idea that contact between Northern and Southern dialects played a causal role. **Fischer et al. (2000)** is an overview of the state of the art concerning generative studies of the historical syntax of English. **Görlach (1991)** is a very useful general overview of Early Modern English. **Gray (1985)** is a collection of Middle English texts, with commentary.

Germanic syntax

Holmberg (1986) is a study of word order and case in the Scandinavian languages, and features the original statement of Holmberg's generalization: the object moves only if the verb does. **Holmberg (1999)** restates the generalization in the light of new data and theoretical developments. **Roberts (1995)** shows that Holmberg's generalization, in its earlier form, held in Early Modern English for as long as verb-movement was still found. **Rögvaldsson (1996)** is an early study of word-order change in the history of Icelandic, which uses the 'grammars-in-competition' approach. **Hróarsdóttir (1996; 1999; 2000)** together represent the most detailed studies of word-order change in the history of Icelandic to date. **Penner and Bader**

(1995) provide a series of highly detailed analyses of V2 and related phenomena in Swiss German. They give evidence that in certain cases V2 affects the interpretation of the clause. **Müller (2004a)** proposes a novel analysis of V2, in which he argues that, rather than involving the combination of verb-movement with fronting of an XP, it involves fronting of a remnant category which contains just the finite verb and an XP. **Bobaljik and Jonas (1996)** propose an early minimalist analysis of Germanic transitive expletive constructions.

Older Indo-European languages

W. Lehmann (1993) is an in-depth study and a personal view of the state of the art in Indo-European studies at the time. **A. Taylor (1994)** is a study of the change from OV order in Homeric Greek to VO order in Classical Greek, using the grammars-in-competition approach. **Vincent (1988)** is an overview of the structure of Latin, with a particularly careful and interesting analysis of the clausal complementation system of that language. We will look at this much more detail in §2.4.

Other important works mentioned in this chapter

Matthews (2001) is an overview of the development of structural linguistics, both in Europe and the USA, in the twentieth century. **Roberts and Roussou (2003)** present a formal account of grammaticalization, treating it as categorial reanalysis driven by change in properties of functional heads. We will say more about grammaticalization in §2.2. **Labov (1972)** is a collection of classic articles on sociolinguistics from the 1960s and 1970s. Most relevant to our concerns is his study of negative concord in Afro-American Vernacular English (AAVE). We return to this in more detail in §4.2.2. **Bhatt and Pancheva (2004)** is a recent analysis of comparative constructions in minimalist terms, with a novel approach to the relation between the comparative morpheme and the comparative clause. **Bianchi (1999)** is a thorough investigation of the syntax of relative clauses and related constructions, developing a ‘head-raising’ analysis (in which the relative head undergoes a kind of wh-movement from its position inside the relative) in terms of Kayne’s antisymmetric model of phrase structure.

This page intentionally left blank

2

Types of syntactic change

Introduction	121	2.5. Word-order change: OV > VO in English	175
2.1. Reanalysis	122	2.6. Conclusion to Chapter 2	198
2.2. Grammaticalization	141	Further reading	198
2.3. Argument structure	149		
2.4. Changes in complementation	161		

Introduction

Unlike the previous chapter, here the discussion will focus entirely on diachronic questions. Now that the notion of parametric change has been justified on empirical grounds, my goal here is to discuss a number of different kinds of syntactic change, and to show how the notion of parametric change can account for them. Thus my goal is to illustrate the power and utility of the parametric approach to syntactic change. The goal of the last chapter was to show that parametric variation was operative in the diachronic domain, i.e. that at least *some* examples of syntactic change can be analysed as parameter change. Here I want to show that *all* the major kinds of syntactic change involve parameter change. Thus the notion of parameter is not merely useful, it is pervasive; in fact, I wish to maintain that it is the principal explanatory mechanism in diachronic syntax. This is not to imply that non-parametric change does not exist; it does, and we will see an example of it in §2.3.

In Chapter 1 I worked with a rather rough-and-ready notion of parameter. In fact, I offered no general definition, still less a formal or technical one, of what a parameter might be. I will continue in this vein in this chapter, although the notion will be made slightly more precise. In Chapter 3 I will offer a more formal characterization. For present purposes, it is sufficient to work with a rather general and informal notion: a parameter is a dimension along which grammatical systems may vary.

In §2.1 I look at reanalysis, which has frequently been considered a mechanism of syntactic change (see Andersen (1973); Lightfoot (1979); Harris and Campbell (1995); Roberts (1993a)). Here I will try to show how, properly defined, reanalysis is forced by parameter change. §2.2 deals with grammaticalization, the development of new grammatical elements from other grammatical elements or ‘full’ lexical items. The phenomena will be discussed and illustrated, and the formal analysis summarized, following the main ideas put forward by Roberts and Roussou (1999; 2003). In §2.3 I turn to changes in argument structure; perhaps the best known case of this kind of change is the development of psychological predicates in the history of English. This will be summarized, and a partially parametric analysis discussed, developing and updating certain ideas in Lightfoot (1981); Fischer and van der Leek (1983); Kayne (1984); and in particular Allen (1995). In §2.4 I discuss changes in clausal complementation, taking the very well-known and extensive changes that can be observed in the development from Latin to Romance as the principal example (see Vincent (1988: 65–73) for a summary of these). Again, I will propose that these changes represent changes in parameter values. Finally, §2.5 picks up the discussion of word-order change from Chapter 1, and discusses word-order change in the history of English in some detail; this leads us to a more refined approach to the variation in word order than was described in §1.6.

2.1. Reanalysis

2.1.1. *The nature of reanalysis*

Harris and Campbell (1995: 50, 61) define **reanalysis** as ‘a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any modification of its surface manifestation’, although they add that there can be a surface manifestation in the form of word-order

or morphological change, perhaps appearing after the reanalysis of underlying structure has taken place. What I want to show here is that reanalysis is intimately bound up with parameter change. In fact, reanalysis is usually a symptom of a change in the value of a parameter; given the central idea that parameters unify clusters of surface grammatical properties, this implies that a parameter change may manifest itself as a cluster of reanalyses, and a reanalysis is usually one symptom of a parameter change.

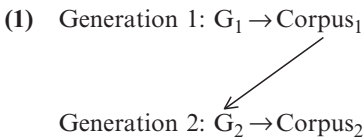
The idea that reanalysis is central to syntactic change is hardly novel. Harris and Campbell (1995: 16) show that it may go back as far as Aristotle and the Arabic grammatical tradition. Later (31–2), they give examples of the concept from the writings of Bopp (1816); Paul (1920); Brugmann (1925); and Wackernagel (1926–8). They also state (30) that reanalysis ‘has been perhaps the single most important factor in modern treatments of syntactic change’. If we can relate reanalysis to parameter change, then, we will clearly be giving parameter change a central role in diachronic syntax.

In a sense, we have little choice other than to relate reanalysis to parameter change, given our general assumptions. Following Harris and Campbell’s definition, reanalysis affects the structural representation associated with a surface string, without altering the string itself. The structural representation, given the assumptions made up to now, is built from three major operations: Merge, Move and Agree. Of these, Merge is the most fundamental operation, since it creates structure: we might think that it is invariant, as in fact was tacitly assumed when this operation was presented in the Introduction. (I return to this point in §2.5.) If so, then it cannot be open to reanalysis. Move and Agree are subject to parametric variation, as we saw in detail in Chapter 1. Hence parameters relating to these operations are what changes when reanalysis takes place. We will see examples of this below.

Beginning, it seems, with Paul (1920) (see Harris and Campbell’s (1995: 31) discussion of his ideas), reanalysis has often been related to child-language acquisition. An important concept here is that of **abductive change**, as put forward, in the context of a discussion of phonological change in Czech, by Andersen (1973). Abduction was distinguished from induction and deduction by the philosopher Charles Sanders Peirce. Deduction proceeds from a law and a case to a result (for example, ‘All men are mortal’ (law); ‘Socrates is a man’ (case), therefore ‘Socrates is mortal’ (result)). Induction proceeds from a case and a result to a law (for example, an immortal being may observe that men (cases) eventually die

(result) and conclude that all men are mortal (law)). Abduction proceeds from a law and a result to a case. Abduction is open to error in a way that induction and deduction are not. With deduction, the case instantiates the law, and so the result must follow. With induction, the result is intrinsically associated with the case, and so the law follows. But abduction cannot follow necessarily: the connection between the case and the results known to follow from the law might be accidental. To take our example, from the statement 'x is mortal' (result) and the law that all men are mortal, one cannot conclude that x is human (case). It is easy to see that x could be a mortal non-human.

In part because of its logically flawed nature, the notion of abduction gives us a useful way of thinking about reanalysis in language acquisition. Following Andersen (1973: 767), we can schematize abductive change as follows:



Here, 'Corpus' refers to a body of sentences produced by speakers. This is called an 'Output' by Andersen, and, in work on **learnability** it is called a 'text'; see the introduction to learnability theory in Bertolo (2001). 'G(rammar)' refers to an instantiation of UG with parameters set. Generation 1 (which we can think of, somewhat simplistically, as the 'parental' generation, the term 'generation' being intended in its everyday sense) has grammar G_1 which underlies Corpus_1 . Generation 2's grammar (simplistically, the 'children's grammar'),¹ G_2 , derives from Corpus_1 and Universal Grammar, given the assumptions about language acquisition we have adopted here (which were summarized in §1.1). The notion of abduction comes in here, since we can think of UG as the law, and Corpus_1 as the result: the child then abduces the case, i.e. a particular grammar. But, as illustrated above, the child may make an error of abduction, and, as it were, mistake a similar case (G_2) for the actual case, G_1 . The important thing about language acquisition that the schema in (1) brings out is that there is no direct link between G_1 and G_2 . This is because, in the last analysis, grammars are mental entities and it is impossible to have direct access to the contents of another mind. Grammars are only transmitted from one generation to the

¹ In §4.2 we will see a reason to modify this simplistic terminology.

next via corpora, and corpora may give rise to errors of abduction. Still putting things rather simplistically, the possibility arises of ‘mismatches’ between G_1 and G_2 as a consequence of the way in which grammars are transmitted. The general view has been that reanalyses are just such mismatches. This has been widely regarded as the basic factor underlying change (see the discussion in Harris and Campbell (1995: 30–2, 61ff., and the references given there). To quote Kroch (2000: 699): ‘[I]anguage change is by definition a failure in the transmission across time of linguistic features’.

If syntactic change centrally involves reanalysis and if reanalyses are mismatches, and if reanalysis is symptomatic of parameter change, then it follows that parametric change is the basic factor underlying syntactic change. Moreover, if reanalysis is driven by abduction in language acquisition, then so is parameter change. So we arrive at one of the main ideas we will explore in this book (mainly in the next chapter): that parametric change is driven by language acquisition. As already stated, this idea is not new: it seems to have first been put forward by Hermann Paul, and has been argued for recently most notably by Lightfoot (1979; 1991; 1999).

This scenario for abductive change naturally raises two fundamental questions: What are ‘mismatches’? and How can mismatches arise? Let us assume for the moment that ‘mismatches’ are reanalyses in exactly the sense defined by Harris and Campbell as given above, and that these must be linked to a parameter change; at the abstract level, mismatches must be connected to parametric options associated with the operations Move and Agree. Then we can see that Generation 2 may abduce some difference in underlying structure for some part of Corpus_1 as compared to Generation 1, and this may have some effect (in morphology or word order, as Harris and Campbell suggest) on Corpus_2 ; these effects are the overt signs of the parameter change.

Putting things this way brings out the problems with this approach. There are two principal problems, which we can call the Regress Problem and the Chicken-and-Egg Problem.² The Regress Problem can be put as

² Croft (2003: 247) also refers to a ‘Chicken-and-Egg Problem’ in diachronic syntax. But his problem is different from the one I discuss below. Croft’s problem is that reconstructed changes may be used to support hypotheses about typological change, while a postulated typological change may be supported by a reconstructed change. As Croft says: ‘[t]his appears to be a vicious circle.’ As we will see below, however, the Chicken-and-Egg Problem for us relates to distinguishing causes and effects of change, which is a different matter.

follows: an innovation in Corpus_2 may be ascribable to a mismatch in G_2 (compared to G_1), but it must have been triggered by something in Corpus_1 – otherwise where did it come from? But if Corpus_1 could trigger this, then how could G_1 produce this property without itself having the innovative property? To quote Kroch (2000: 699–700) again:

Since, in an instance of syntactic change, the feature that learners fail to acquire is learnable in principle, having been part of the grammar of the language in the immediate past, the cause of the change must lie either in some change, perhaps subtle, in the character of the evidence available to the learner or in some difference in the learner, for example in the learner's age at acquisition, as in the case of change induced through second-language acquisition by adults in situations of language contact.

Here Kroch illustrates the problem and the only possible solutions: either Corpus_1 is subtly changed so that G_2 is more readily abducted from it than G_1 , or some external factor such as language contact is at work. There is no doubt that language contact plays an important role in many syntactic changes, and that it can provide a straightforward solution to the Regress Problem. This will be the subject matter of Chapter 5. But it seems that not all changes can be explained through contact, and where contact is not a causal factor, subtle changes in Corpus_1 seem to offer the only mode of explanation for change. These subtle changes may be caused by some extrasyntactic, but still intralinguistic, factor such as phonological or morphological change; we will see examples of this below.

If some change in Corpus_1 is responsible for reanalysis but is not itself the reanalysis, we face the Chicken-and-Egg Problem. If we observe two correlated changes, how can we know which caused the other? To put it another way, we might want to say that two innovations in Corpus_2 are due to a single mismatch in G_2 caused perhaps by a single feature of Corpus_1 . This will solve the Regress Problem along the lines just sketched, for one of the innovations. But if Corpus_1 shows the two innovations, how do we know which is playing the causal role? How do we know which innovation is a cause and which an effect of the reanalysis? And, for whichever one we call the cause, we still have the Regress Problem. This problem can be observed in two different treatments of the causal role of reanalysis. On the one hand, Lightfoot (1979) proposes a series of different changes leading to accumulated opacity in the grammar, ultimately causing a reanalysis (we will see an example of this directly); on this view, the prior changes are not explained and are subject to the Regress Problem, although the reanalysis is explained. On the other hand, Timberlake (1977) and Harris and Campbell

(1995) propose that reanalysis causes a group of unrelated changes; this approach explains the changes but not the reanalysis (see Harris and Campbell (1995: 77)). Of course we are always free to assert that there is no causal relation between the two innovations, but in doing this we flout Occam's razor (by having more entities, i.e. underlying changes, than necessary) and have the Regress Problem twice over.³

2.1.2. *The Transparency Principle*

Lightfoot's (1979) Transparency Principle offered a way of dealing with these problems, as can be seen from his discussion of the development of English modal auxiliaries (*can, must, may, will, shall, ought*). Lightfoot argues that several changes affecting these items took place together in the sixteenth century.⁴ These include the loss of the ability to take direct objects (or indeed any kind of complement other than an apparently bare VP, with *ought* a consistent exception in requiring a *to*-complement), and the loss of non-finite forms. (2) illustrates an early example of *will* with a direct object, and an example of an infinitival modal (*konne*, corresponding to NE *can*):

- (2) a. Wultu kastles and kinedomes?
 Wilt thou castles and kingdoms?
 (c. 1225, Anon; Visser (1963–73, §549))
- b. I shall not konne answer.
 I shall not can answer
 (1386, Chaucer; Roberts (1985: 22))

Moreover, after the loss of V-to-T movement (the change in parameter B discussed in §1.3.2), modals diverged syntactically from all the other verbs

³ Harris and Campbell (1995: 40–4) criticize Lightfoot's (1991: 166ff.) discussion of the differences between parametric changes and other kinds of changes in part because it does not solve the Chicken-and-Egg Problem. The criticisms are partly justified, but they apply to any approach involving reanalysis, as Harris and Campbell (77) acknowledge.

⁴ Many authors have pointed out that Lightfoot's chronology seems to be incorrect, in that it is not clear that all these changes took place at the same time; see in particular Warner (1983; 1993). However, I present the development approximately as Lightfoot did, since it illustrates the general point regarding transparency and reanalysis that I wish to make here.

of English (except the aspectual auxiliaries *have* and *be*, and dummy *do*) in that they retained the earlier pattern of negation and inversion syntax, i.e. they precede clausal negation and are inverted over the subject in main-clause interrogatives. This is of course still the case in present-day English:

- (3) a. I cannot speak Chinese.
b. Can you speak Chinese?

This can be accounted for, consistently with the idea that parameter B changed value in the sixteenth century, if we assume that by the time this parameter changed, modals were merged in T rather than V. Hence, once the V-to-T parameter changed, the syntactic differences between modals and main verbs in negation and inversion emerge since main verbs no longer move to T, while modals are merged there. So we have the NE situation in which modals have ‘T syntax’ and main verbs have ‘V syntax’.

According to Lightfoot (1979), the creation of a new class of modal auxiliaries was due to the accumulation of exception features – morphological, semantic, and syntactic – on the modal verbs, which made them ‘opaque’ as main verbs. The morphological exception feature was that the modals, by the sixteenth century, were the only surviving members of the class of OE ‘preterit-present’ verbs. These verbs are characterized by having ‘a strong past tense with present meaning . . . and a new weak past tense’ (Mitchell and Robinson 1992: 52). By late ME, the consequence of this was that these were the only verbs in the language to lack a 3sg ending in the present tense (*-(e)s* or *-(e)th*); in a language with as impoverished an inflectional system as English, it is reasonable to suppose that this is a highly irregular feature. The semantic ‘irregularity’ of these verbs was their modal meaning, and in particular their ability to form a periphrastic substitute for the moribund subjunctive inflections. In virtue of their meaning, the usual form–meaning correlation between preterit morphology and past time did not always hold (for example, in *I should do it tomorrow*). One syntactic irregularity may have been that, with the glaring and unexplained exception of *ought*, the modals never took *to*-infinitives as their complement, although this was established as the main form of non-finite sentential complementation by the end of the ME period (Los 1998; Fischer *et al.* 2000: 211ff.); Lightfoot (1979: 101–9) is the original presentation of these and other opacity-inducing factors.

So, Lightfoot’s claim is that the Transparency Principle forced the modals to change category once this opacity became too great. This approach

has two notable advantages. First, it narrows down the Regress Problem; as long as we know how much opacity can be tolerated and what the nature of opacity really is, we can know at which point Corpus₁ will have sufficient exception features to cause Generation 2 to abduce G₂ rather than G₁. More precisely, suppose the Transparency Principle states that a certain structure can only be acquired if it requires the postulation of less than n exception features. G₁ is acquired on this basis, but something in Corpus₁ must be abduced as a further exception feature, making G₁ unlearnable for Generation 2, and hence triggering reanalysis. We can see that the Regress Problem still appears in that there is some feature of Corpus₁ which must be an exception for Generation 2 but not for Generation 1. Similarly, a characterization of exception features would also solve the Chicken-and-Egg Problem; otherwise this arises in connection with exactly the same feature of Corpus₁, which is what is really driving the reanalysis. Nevertheless, the merit of the Transparency Principle is that it forces us to say that reanalysis is caused by one exception feature too many.

The problems with the Transparency Principle also emerge from this discussion. The most fundamental of these is that there is no definition of transparency or its converse, opacity. Without these notions, it is clear that the potential advantages relative to the Regress Problem or the Chicken-and-Egg Problem are not realizable. Unfortunately, Lightfoot (1979) offered no such definitions, and neither have any arisen in more recent work by Lightfoot or others. So we must conclude that the Transparency Principle does not offer true solutions to the Regress Problem and the Chicken-and-Egg Problem.

2.1.3. *Phonology and reanalysis*

One way to tackle both the Regress Problem and the Chicken-and-Egg Problem is to attribute the crucial factor leading to reanalysis to another part of the grammar, for example, phonology or morphology. An example where phonology plays a role is the development of the question particle *ti* in Colloquial French (see Harris (1978); Bennett (1979); Roberts (1993a: 222–4); Harris and Campbell (1995: 66); a similar development has taken place in the history of Occitan (Wheeler 1988: 272–3) and some varieties of Franco-Provençal Valdôtain (Roberts 1993b: 342ff.)). This element is a reanalysis of the epenthetic consonant /t/ and the 3sg masculine pronoun *il* in inversion contexts, roughly as follows:

- (4) (Jean) a-t-il fait cela? → Jean a ti fait cela?
 (John) has he done that John has Q done that
 ‘Has John done that?’

This change, which in fact involves the reanalysis of subject-clitic inversion and the loss of pronominal features associated with *il*, depends on the ability to drop word-final /l/ after /i/ in colloquial French. The effects of it can be seen where the preverbal subject is not 3sg masculine, as in:

- (5) a. Elle t'écrit ti souvent?
 she you-writes Q often
 ‘Does she write to you often?’
 b. On t'a ti demandé ton adresse?
 one you-has Q asked your address
 ‘Have you been asked for your address?’

Also, since the ‘complex inversion’ construction from which *ti* was reanalysed could not have an initial subject clitic (**Il habite-t-il Lyon?* ‘he lives-he in Lyon’ = ‘Does he live in Lyon?’; see Kayne (1983); Rizzi and Roberts (1989) on this), the existence of examples with *il* in subject position is a further indication of this reanalysis:

- (6) Il habite ti Lyon?
 he lives Q Lyon
 ‘Does he live in Lyon?’

For this reanalysis to take place, it suffices that Generation 1 produced an inversion structure containing epenthetic /t/ and the pronoun *il*, with a low-level phonological rule deleting word-final /l/. This gives rise to a surface string containing the phonological sequence /ti/. By treating /ti/ as a Q-marker, Generation 2 can analyse this string as containing no postverbal subject clitic (in the complex inversion construction, a preverbal subject is present in any case), no /t/-epenthesis, and no /l/-deletion. Syntactic opacity may play no role here; rather it is the indeterminacy of the earlier form which makes it subject to reanalysis (although I will return to this point directly). Both the Regress Problem and the Chicken-and-Egg Problem are solved by appealing to the idea that the crucial causal factor was the deletion of word-final /l/.⁵

⁵ Actually the problems are solved for syntax, but they may be shifted to the phonology. If final /l/-deletion is a productive option, why is it not postulated by Generation 2 in this case? Again, we do not fully understand why Generation 1 tolerates the earlier grammar and why Generation 2 innovates. See note 6.

Roberts (1993a: 155ff.) puts forward a general notion of Diachronic Reanalysis (DR) which is operative here. The reanalysis relating the two constructions in (4) is given in (7); (see Roberts (1993a: 222), although the structures proposed here are simpler in various respects) the *(-t-)* in (7a) is presumably not present in the syntactic structure, being an epenthetic consonant:⁶

- (7) a. [_{CP} Jean [_C [_T a] C] [_{TP} (-t-) il [_{VP} fait cela]]] >
 b. [_{TP} Jean [_T a] ti [_{VP} fait cela]]

This change can be dated to the early seventeenth century (Roberts 1993a: 223–4). According to Roberts (155ff.), both structural ambiguity and structural simplicity are preconditions for a DR of this type in that (7b) is clearly a simpler structure than (7a). I will discuss various ways of characterizing structural simplicity in §3.4. Hence opacity does in fact play a role, in the guise of structural simplicity; the idea is that reanalysis is motivated by a general preference on the part of language acquirers to assign the simplest possible structural representations to the strings they hear (as part of Corpus₁). I will henceforth refer to this as the ‘simplicity preference’.

Moreover, DR of the type illustrated in (7) is associated with parameter change. DRs are seen as the symptoms of parameter change. Here, the development of *ti* is associated with the loss of subject–clitic inversion in main-clause yes–no questions; to the extent that inversion involves T-to-C movement of the kind described in §1.3.1, and depends on what we might call the ability of the relevant type of C to trigger movement (see §2.5 for more on this), it is a property subject to parametric variation. Roberts (159) suggests that ‘the notion DR may . . . prove to be epiphenomenal. All DRs may turn out to be instances of Parametric Change.’

⁶ See Roberts (1993a: 221), and the references given there, on the structural position of *ti*. It is clearly lower than the position of the finite auxiliary, which we are assuming to be T, but external to VP. It is possible that the phonological opacity concerns /t/-epenthesis rather than /l/-deletion. No morphological or phonological operation equivalent to /t/-epenthesis is found elsewhere in French, while final-consonant deletion is rife on most analyses (see Dell (1985); Tranel (1981); Pagliano (2003)). Moreover, the opacity of /t/-epenthesis would carry over to the many varieties (including Quebec French, as well as varieties of Occitan and Franco-Provençal) where the question particle appears to have arisen from the 2sg pronoun *tu*. I leave these complex and interesting questions aside here.

The approach to reanalysis which regards it as caused both by simplicity and ambiguity gives rise to an interesting angle on the Regress Problem: assuming that language acquirers (Generation 2 in (1)) will always prefer the simplest possible representation of the strings of Corpus₁, we have to look for what *prevented* the simpler analysis for Generation 1. In this case, we take this to be phonology; Generation 1 has an underlying /l/ in *il*, which is deleted by the /l/-deletion rule (see notes 5 and 6 for some provisos to this). Similarly, the Chicken-and-Egg Problem may be reduced to phonology; presumably some change in the underlying phonological form led Generation 2 to abandon the underlying /l/ in *il*, with the reanalysis as a direct consequence. From the point of view of syntax, then, the problems are solved, although they may resurface in accounting for the relevant phonological changes.

Lightfoot (1999: 216–17) critiques DRs for having no really useful role to play in an account of language change. Strictly speaking, this may be true; we have already seen that Roberts (1993a) suggests DRs may be epiphenomenal, and we are following that suggestion here. Lightfoot correctly states that DRs are to be regarded as relating grammatical representations of subsequent generations, but incorrectly points out that ‘they occur where grammatical shifts have already taken place’ (217). In fact, DRs are intended as an indication of how a potentially ambiguous string had one analysis at one period (Generation 1) and another at a later period (Generation 2). Their utility lies in bringing out the alleged role of simplicity and ambiguity in driving reanalysis.

We have seen that Lightfoot (1979) regards opacity as the principal cause of reanalysis, although he also mentions ambiguity (1979: 351). Timberlake (1977) and Harris and Campbell (1995: 70ff.) consider reanalysis to be a consequence of ambiguity. Finally, Roberts (1993a) regards reanalysis as driven by both factors, assuming that the preference for simplicity can be seen in terms of opacity of the earlier structure.

2.1.4. *Expressing parameters*

If we are to view reanalysis as always accompanying parameter change, i.e. as the structural manifestation of the change in the value of at least one parameter, then we have to consider how Corpus₁ in (1) succeeds or fails in triggering different values for a given parameter, i.e. in leading language

acquirers to set a given parameter to a given value. Lightfoot (1999) tackles this question, following Dresher (1999), by introducing the notion of **cue** for a parameter. See also Lightfoot (2006: 82ff.), where a number of examples of cues are given; the loss of V-to-T movement in ENE and the development of the modals and *do* are also discussed there (90–100). In a similar vein, Clark and Roberts (1993) introduced the concept of parameter expression. This can be defined as follows (this definition is from Roberts and Roussou (2003: 15)):

- (8) A substring of the input text *S* expresses a parameter p_i just in case a grammar must have p_i set to a definite value in order to assign a well-formed representation to *S*.

To give a simple example, a sentence like (9) (repeated from §1.2.1), expresses the positive value of the null-subject parameter, since this parameter must be given the positive value in order for the sentence to be grammatical:

- (9) Parla italiano.
S/he speaks Italian.

The notion of ‘trigger’ (or, equivalently, cue) can be defined in terms of parameter expression, as follows:

- (10) A substring of the input text *S* is a trigger for parameter p_i if *S* expresses p_i .

Thus (9) is a trigger for (the positive value of) the null-subject parameter. Clearly, for Generation 2 to converge on the same grammar as Generation 1 in the scenario in (1), Corpus_1 must express all the parameters of UG.

We can begin to connect P-expression to reanalysis by introducing the following notions (again, originally from Clark and Roberts (1993), but slightly reformulated here):

- (11) a. P-ambiguity:
A substring of the input text *S* is strongly P-ambiguous with respect to a parameter p_i just in case a grammar can have p_i set to either value and assign a well-formed representation to *S*.
- b. A strongly P-ambiguous string may express either value of p_i and therefore trigger either value of p_i .
- c. A weakly P-ambiguous string expresses neither value of p_i and therefore triggers neither value of p_i .

Strong P-ambiguity is arguably linked to reanalysis. We might suppose that reanalysis takes place given a class of strongly ambiguous strings in

relation to a particular parameter in a given corpus, and where a simpler representation is associated with one value rather than the other. In the example involving French interrogative *ti* given above, the relevant strings are rendered P-ambiguous with respect to subject-clitic inversion (T-to-C movement; the ‘residual’ version of the V2 parameter of §1.3.1.2) by the phonological option of /l/-deletion or selection of an underlying form lacking final /l/. Since the reanalysed structure is simpler than the earlier one (see (7) and the following discussion, as well as notes 5 and 6 on /t/-epenthesis), this is the preferred structure. So P-ambiguity and the simplicity preference are what drives reanalyses, seen as surface manifestations of parameter change.

We can give a more extended example of how this approach works with the loss of V-to-T movement (parameter B of Chapter 1) in ENE. As we saw in Chapter 1, examples like the following (repeated from (74)) indicate that V moves to T at this period:

- (12) a. if I **gave not** this accompt to you
 ‘if I didn’t give this account to you’
 (c1557: J. Cheke, Letter to Hoby; Görlach 1991: 223; Roberts 1999: 290)
- b. The Turkes . . . **made anone redy** a grete ordonnaunce
 ‘The Turkes . . . soon prepared a great ordnance.’
 (c1482: Kaye, *The Delectable Newsse of the Glorious Victorye of the Rhodyans agaynest the Turkes*; Gray 1985: 23; Roberts 1993a: 253)

In terms of the notion of P-expression introduced above, we can say that examples like this express the positive value of the V-to-T parameter. On the other hand, at that time as at this, many very simple sentences, which must have been extremely prominent in the trigger experience, were P-ambiguous. A simple sentence such as *John walks* expresses either value of V-to-T, as illustrated by the two possible structures in (13), and is strongly P-ambiguous:

- (13) a. John [_T walks] [_{VP} . . . (walks) . . .]
 b. John T [_{VP} walks]

Furthermore, following the change in status of the modal auxiliaries and *do* (which appears to have taken place slightly earlier in the ENE period than the loss of V-to-T; see Roberts (1993a: 310ff.); Warner (1997: 382–3); and below), any simple sentence containing a finite auxiliary was weakly P-ambiguous regarding the V-to-T parameter, assuming the auxiliary was in T:

- (14) a. I may not speak.
b. I do not speak.

So we see that there was much P-ambiguity regarding the V-to-T parameter in sixteenth-century English. However, this ambiguity existed, albeit in a slightly different form, in ME too. (14) was strongly P-ambiguous prior to reanalysis of modals and *do* as auxiliaries, since these elements were at that stage main verbs. In this connection, the fact that dummy *do* became an auxiliary at the same time as the modals (see Denison (1985) and Roberts (1993a: 295)) played a very important role. This is the case because in the sixteenth century *do* could appear in positive declaratives, as shown in (15) (both examples are from Shakespeare's *Richard III*, discussed in Barber (1976: 164)):

- (15) a. Where eyes did once inhabit.
b. Thou didst receive the sacrament.

In fact, *do* could seemingly appear in any context, except where a modal was present. Thus, *do* was always available as an alternative to verb-movement. In particular, this meant that there was always a non-V-movement alternative to constructions like (12), which otherwise expressed the positive value of the V-to-T parameter.

We still have to ask why it is that the P-ambiguity of examples like (13) and (14) only became crucial in the sixteenth century. In other words, what prevented this P-ambiguity from leading to a reanalysis of V-to-T movement structures prior to this time? One answer has to do with morphology. Southern varieties of English lost a large part of their verbal agreement morphology in the latter part of the 15th century.⁷ For example, Gray (1985: 495ff.) gives shown in Table 2.1 agreement paradigms for the present

⁷ Northern varieties, notably including Older Scots (spoken and written in the Kingdom of Scotland in the fifteenth and sixteenth centuries – see Derrick McClure (1994)) had rather different paradigms from OE. By the sixteenth century, these paradigms were apparently invariant, although they were already subject to what may have been a precursor of the modern Northern Pronoun Rule, in that the agreement endings disappeared in certain persons where the subject was non-pronominal; see Roberts (1993a: 265ff.) and the references given there; C. Jones (1997) on Scots varieties from a synchronic and diachronic point of view; Henry (1995) on the variant of the Northern Pronoun Rule found in present-day Belfast English (a variety which derives from Older Scots); and Jonas (2002) on the present-day Shetland dialect of English.

Table 2.1 Verbal agreement inflection in Middle English

1400	1500
cast-(e)	cast
cast-est	cast-est
cast-eth	cast-eth
cast-e(n)	cast-(e)
cast-e(n)	cast-(e)
cast-e(n)	cast-(e)

tense of the verb *cast* in East Midlands English at the beginning and end of the fifteenth century.

Shortly after 1500, what remained of the plural agreement marking was lost. This development is striking in that many authors have observed a correlation between the ‘richness’ of verbal agreement inflection and the positive value of the V-to-T parameter, notably in a range of Scandinavian languages and dialects. Vikner (1997: 201) sums up the relation as follows:

- (16) An SVO language has V-to-T movement if and only if person morphology is found in all tenses.

This generalization (and its precursors; see the very thorough discussion of these in Vikner (1997)) has been criticized for being empirically too strong. There appear to be a number of varieties in which verbal inflection has disappeared or nearly disappeared, but which nevertheless continue to show V-to-T movement. Two well-known cases are the Kronoby dialect of Swedish (spoken in Finland) and the Norwegian dialect of Tromsø:

- (17) a. He va bra et an **tsöfft** **int** bootsen.
 it was good that he bought not book-the
 ‘It was good that he didn’t buy the book.’
 (Kronoby; Platzack and Holmberg 1989: 74)
- b. Vi va’ bare tre stökka før det at han Nielsen **kom** **ikkje**.
 we were just three pieces for it that he Nielsen came not
 ‘There were only three of us because Nielsen didn’t come.’
 (Tromsø Norwegian; cf. Vikner (1997: note 19, 211))

Here we see that the finite verb in the embedded clause precedes negation. These examples are therefore equivalent to sixteenth-century English examples like (14a), and are taken to indicate that V moves to T in these

varieties. However, these varieties have no subject-verb agreement at all (like the standard Mainland Scandinavian languages, which lack V-to-T movement). As Thráinsson (2003) points out, this indicates that a biconditional statement of the type in (15) cannot be right (see also Roberts (1993a: 267); Bobaljik (2002); Alexiadou and Fanselow (2002) on this and related matters).

Instead, following Roberts (1999: 292), we may think that morphological paradigms of certain types may express parameter values. For example, let us restate Vikner's generalization as follows:

- (18) If (finite) V is marked with person agreement in all simple tenses, this expresses a positive value for the V-to-T parameter.

Example (18) differs from Vikner's generalization in two important ways. First, it is a statement about the expression of a parameter, and thus ultimately about the trigger experience, rather than being a statement about something internal to UG. In other words, it represents 'a choice from among the surface cues from among the limited set of possibilities provided by Universal Grammar' in the words of Anderson (2002: 273), who criticizes approaches of the type put forward by Vikner in which morphology determines syntax. Second, it is a one-way implication; it allows for languages with a positive value of the V-to-T parameter but without verbal agreement inflection, just as has been observed in varieties such as Kronoby Swedish, and Tromsø Norwegian. Thráinsson (2003: 154) similarly proposes a one-way implicational relation between V-to-T movement and the relevant verbal agreement inflection. See also Bobaljik and Thráinsson (1998).

Example (18) applies to the past, too. So, as mentioned in note 7, Middle Scots had a seemingly invariant verbal agreement paradigm (with the complication mentioned there) and yet allowed V-to-T movement, as examples like the following show:

- (19) Quhy **sing ye nocht**, for schame!
 why sing you not, for shame
 (Anon. *The Unicornis Tale*; Gray 1985: 158; Roberts 1993a: 266)

So we can conclude that, prior to 1500 or shortly afterwards, verbal agreement inflection in Southern varieties of English expressed a positive value for the V-to-T parameter. Interestingly, there was a delay between the loss of agreement marking and the loss of V-to-T movement, in that verbal inflection is lost approximately seventy-five years before V-to-T movement.

Thráinsson (2003: 184–5) shows that, as far as can be ascertained, the same is true for the Mainland Scandinavian languages – notably standard Swedish and Danish – which have historically lost V-to-T movement and verbal agreement marking. All this is consistent with the one-way implicational statement in (18).

In fact, Warner (1997: 382–3) divides the chronology for the loss of V-to-T in English into four periods. In Period 1 (up to ca. 1500) T attracts V, due to its agreement morphology, as we have seen. In Period 2 (1500–roughly 1700) T loses the attraction property and variation ensues as data like (12) triggers V-to-T, but the evidence of modals and *do* in T does not favour this, being weakly P-ambiguous in relation to this parameter. In Period 3 (ca. 1700–50) V-to-T movement is no longer found, but there are lexical exceptions (mainly *know* and *doubt*) which continue to show the older pattern. Finally, in Period 4, from 1750 on, V-to-T movement of main verbs no longer occurs.

The shift from Period 1 to Period 2 is the crucial one in the present context. If this line of reasoning is correct, the loss of morphological expression of the V-to-T parameter created the strong P-ambiguity needed for a reanalysis of the following kind:

- (20) [_{TP} John [_T walk-eth] ... [_{VP} ... (V) ...]] >
 [_{TP} John T ... [_{VP} ... [_v walks]]]

Following Warner, this reanalysis led to variation for a period, but favoured the innovative, structurally most economical grammar. The reanalysis manifests a change in the V-to-T parameter, which, as we saw in §1.3.1, is associated with a cluster of properties: main-verb inversion in questions, *V-adverb-object* order, *V-not* order, and possibly also pronominal object shift and transitive expletive constructions.

Postulating that the morphological expression of the parameter played a crucial role in preventing the earlier reanalysis effectively deals with the Regress Problem in this instance. Moreover, the relative chronology of the loss of verbal agreement morphology and the slightly later change in the parameter gives us a way of dealing with the Chicken-and-Egg Problem, assuming that the relative chronology indicates the causal relation. Nevertheless, we can ask whether these are really principled and general answers to the problems. In particular, this solution to the Regress Problem shifts it to morphology, rather as the answer proposed in the case of the development of French *ti* sketched above shifted it to the phonology.

Similarly, although the relative chronology gives us a clear indication of the causal relation between the loss of agreement morphology and the change in the V-to-T parameter, it raises the tricky question of the nature of the interim grammar: here the morphological expression of the V-to-T parameter is lost, and yet the positive value appears to remain for a generation or two at least. If Warner is right in saying that there was variation at this period, we may be able to appeal to mechanisms of ongoing change of the kind we will discuss in more detail in §4.1 and §4.2, in particular a version of Kroch's (1989) notion of competing grammars. But it is important to see that what we have just sketched above, while clearly indicating the relations among P-expression, reanalysis and parameter change, undoubtedly leaves important questions open.⁸

One currently spoken Scandinavian language appears to be undergoing the loss of V-to-T movement at present: Faroese has verbal morphology which may be compatible with V-to-T movement according to (17), although the actual incidence of V-to-T movement has been the subject of some controversy. Thráinsson (2003) gives an up-to-date survey of what has been said about Faroese in the recent syntactic literature. Thráinsson concludes that morphological reduction has led to variation among dialects, registers, and age groups regarding the incidence of V-to-T movement. In this respect, the situation is not unlike the one Warner suggested for Period 2 of ENE described above. Moreover, as we shall see in §4.2, we expect variation of this type as a change is ongoing; this lends support to the competing-grammars idea. Contemporary developments in Faroese may well be able to tell us a lot about what happened in sixteenth-century English.

⁸ One of these is the technical question of why morphological marking of agreement on V should be associated with attraction by T. In terms of the theory assumed here, basically that of Chomsky (2000; 2001), T and V Agree for tense features, in that T has interpretable tense features and the morphology on V encodes uninterpretable tense features. T has uninterpretable agreement features, but these can only be valued against DPs. Hence it is unclear why V should be attracted to T (i.e. why we should have Move and Agree holding between V and T) just when V has rich morphological marking of agreement. I will leave this technical point open here.

2.1.5. *Reanalysis and the poverty of the stimulus*

One final point before we leave the discussion of reanalysis. It may seem at first sight that the scenario for abductive reanalysis that we have described is actually inconsistent with the argument from the poverty of the stimulus, in that abductive reanalysis is precisely a case where children do not necessarily converge on the grammar underlying their trigger experience. This point becomes clear if we reconsider part of the quotation from Hauser, Chomsky, and Fitch (2002) given in §1.1:

A child is exposed to only a small proportion of the possible sentences in its language, thus limiting its database for constructing a more general version of that language in its own mind/brain. This point has logical implications for any system that attempts to acquire a natural language on the basis of limited data. It is immediately obvious that given a finite array of data, there are infinitely many theories consistent with it but inconsistent with one another. In the present case, there are in principle infinitely many target systems . . . consistent with the data of experience, and unless the search space and acquisition mechanisms are constrained, selection among them is impossible . . .

Under abductive reanalysis, the child does in fact construct for itself a system which is consistent with the data from its experience but which is not exactly consistent with that underlying the trigger experience, as we have seen. But the important thing is that the search space and the acquisition mechanisms are highly constrained, and so reanalyses, although possible and actually attested (if the view of syntactic change being sketched here is correct), do not vary ‘wildly’ over just any imaginable possibilities. Instead, they appear to be of a rather limited type: it has often been observed that syntactic changes fall into fairly well-defined patterns. (See Harris and Campbell (1995, Chapter 2) for an overview of various approaches to syntactic change.) In terms of the particular technical assumptions about syntactic structure we are adopting here, reanalysis only involves functional categories and only affects the operations of Move and Agree (i.e. not Merge) and may well be subject to further constraints. Thus reanalyses reflect the rather limited range of parametric options UG makes available. Furthermore, as we shall see in §4.1 and §4.2, acquirers can discern and reproduce variation and optionality in the primary linguistic data in their internalized grammars. For this reason, studying them may eventually shed light on an important aspect of linguistic theory.

2.1.6. Conclusion

In this section, I have introduced the central concept of reanalysis, largely following Harris and Campbell's (1995) definition. I have suggested that reanalysis is symptomatic of underlying parametric change, and that it results from the abductive nature of language acquisition. I identified two problems which have often been discussed in the literature, which I called the Regress Problem and the Chicken-and-Egg Problem. These problems were discussed in relation to examples of reanalysis from the literature, although they were not resolved in any general way. I will return to these matters in §3.2, where we will see that the Regress Problem falls under the more general logical problem of language change. In §3.3, where we discuss the nature of the trigger for parameter values in more detail, we will come back to the Chicken-and-Egg Problem.

Bearing all this in mind, we turn in the next section to a well-known and highly pervasive type of syntactic change: grammaticalization.

2.2. Grammaticalization

Grammaticalization can be defined as the process by which new grammatical morphemes are created. The term was first coined by Meillet (1912), although, as Harris and Campbell (1995: 19) point out, the notion certainly predates the introduction of the term; see Hopper and Traugott (2003: 19ff.) for a discussion of the history of the nineteenth-century antecedents of the concept. Over the past twenty years or so, grammaticalization has been the focus of much attention in the typological and functional literature on syntactic change. (See in particular C. Lehmann (1986; 1995); Heine and Reh (1984); Heine, Claudi, and Hünnemeyer (1991); Traugott and Heine (1991); Heine *et al.* (1993); Bybee, Perkins, and Pagliuca (1994); Hopper and Traugott (2003); and the compendium of cases of grammaticalization in Heine and Kuteva (2002).) Less attention has been paid to the phenomenon in more formal approaches to syntax, although Roberts and Roussou (1999; 2003); van Kemenade (2000); Wu (2000); Simpson and Wu (2001); Munaro (2005); Tremblay, Dupuis, and Dufresne (2005); and the papers in Batllori *et al.* (2005) are exceptions. Here I will focus on the formal approach to grammaticalization presented in Roberts and Roussou

(1999; 2003), as this clearly illustrates how the phenomenon may be reduced to reanalysis and associated parameter change.

In terms of the kind of theory of syntax being assumed here, in which many grammatical morphemes (complementizers, auxiliaries, determiners, etc.) are seen as exponents of functional categories, the idea that grammaticalization involves the creation of new grammatical morphemes implies that grammaticalization frequently involves the development of new exponents of functional categories. To the extent that functional categories are the locus of parametric change, i.e. able to trigger the cross-linguistically varying properties of Agree and Move, we can see how creating a new exponent of a functional head F may involve creating new parametric properties – triggering of Agree or Move – associated with F.

A frequently discussed example of grammaticalization involves the history of French negation (Jespersen 1917; Foulet 1990; Hock 1991; Déprez 1997; 1999). In §1.4, we saw how a number of what are now ‘n-words’ in Modern French developed from formerly positive expressions: *aucun* (formerly ‘some’, now ‘no’), *rien* (formerly ‘thing’, now ‘nothing’) and *personne* (formerly ‘person’, now ‘no-one’). We also mentioned that a crucial part of this change, the change of clausal *ne*’s Negation feature from an interpretable to an uninterpretable one, may have correlated in the seventeenth century with the development of the *ne . . . pas* pattern as the standard form of clausal negation, with *pas* bearing the interpretable Negation feature from that time on. What we did not discuss there is the origin of *pas*. This word comes from the noun meaning ‘step’, which still exists in contemporary French. It was grammaticalized as a negative marker at the relevant stage in the history of French.

The development of the two-part clausal negation is part of a series of changes first pointed out by Jespersen (1917) which have become known as Jespersen’s Cycle. They can be illustrated for both French and English as follows:

(21) *Stage 1:*

- | | |
|--------|---------------------------------|
| a. OE: | ic ne secge
I neg say |
| b. OF: | jeo ne dis
I neg say |

Stage 2:

- | | |
|---------------------|---|
| a. ME: | I ne seye not
I neg say NEG |
| b. Standard French: | je ne dis pas
I neg says NEG |

Stage 3:

- a. ENE: I say **not**
 b. Colloquial French: je dis **pas**

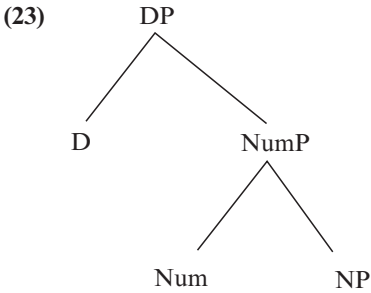
As (21) shows, both French and English illustrate the ‘cyclic’ development from a preverbal negative marker, to a combination of pre- and postverbal marking, to a final stage where only the postverbal marking survives. Stage 3 represents ENE, and, as we have seen, English then developed a rather different pattern of Negation involving the auxiliary *do*; this was linked to the change in the V-to-T parameter which we discussed in the previous section. Like many **grammaticalization cycles**, the changes in the form of Negation illustrated in (21) are not strictly cyclic (see also Hopper and Traugott (2003: 124)), but we can observe an interesting series of apparently related changes. The transition from Stage 2 to Stage 3 may not be of great interest from a syntactic point of view, since it appears to involve just the loss of an unstressed element. On the other hand, the change from Stage 1 to Stage 2 involves the grammaticalization of the negative element. Here I will briefly summarize the development of the French negator *point*, as this is described in Roberts and Roussou (2003: 149ff.). The reason for choosing *point* rather than *pas* is that the development of this element is in certain respects more interesting for our conception of grammaticalization than that of *pas*; recall also that *point* remains an alternative form of sentential Negation alongside *pas*, at least in rather literary varieties of French.

The negator *point* developed from a noun meaning ‘point’; this noun was borrowed into English, and survives in Modern French as a masculine noun. The negator *point*, on the other hand, is not a noun in contemporary French and lacks grammatical gender and other nominal features such as number. This element occupies the same position as *pas*, i.e. it follows a finite verb and precedes a non-finite verb (see §1.3.1); in fact, it is simply a stylistic alternative to *pas* in the relevant registers of French:

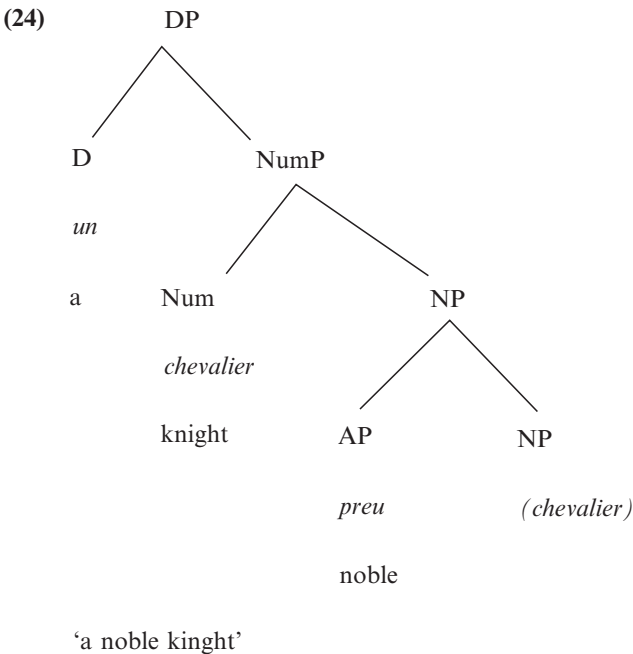
- (22) a. Jean ne mange point de chocolat.
 John neg eats not of chocolate
 ‘John does not eat chocolate.’
 b. ne point embrasser Marie, ...
 Neg not to-kiss Mary ...
 ‘not to kiss Mary ...’

In order to understand the change that converted *point* from a noun into a clausal negator, we need to take a closer look at the internal structure of

DPs and at the development of the French article system. Let us suppose, following Bernstein (1991; 2001); Ritter (1991); Zamparelli (1995), among others, that the structural complement of D is not in fact NP, but rather a further functional category indicating Number, i.e. NumP, as shown in (23) (we briefly mentioned this possibility in §1.5):



Furthermore, we can assume that the usual postnominal position of adnominal APs in both Old and Modern French is a reflex of the fact that nouns in general move to Num, with adnominal APs adjoined to NP (cf. Longobardi (2001: 579–80) and the references given there), as in:



Num is also the position for certain quantifiers, as argued by Zamparelli (1995). In these terms, we can understand the development of a class of n-words in French (including *point*, but also *aucun*, *rien*, *personne*, etc.) as involving the loss of N-to-Num movement for these items and their reanalysis as exponents of Num. This can explain the change in distribution of these elements and the loss of phi-features; after the loss of N-to-Num movement these elements are no longer nouns and so, we may assume, they can no longer enter the relevant Agree relations with Num and D.⁹

How did this change take place? Here we follow Déprez's (1999) analysis. Déprez observes that Modern French DPs almost always require an article. In particular, sentences like (25) are ungrammatical if no article is present:

- (25) Jean a mangé *(des) pommes.
John has eaten (some) apples.

Déprez further observes that this wasn't the case in earlier French. In OF, null Ds are found with singular mass nouns and with bare plurals, much as in English or in other Romance languages (see Longobardi (1994) on the latter point):

- (26) a. Si mengierent **pain** et burent **cervoise**.
so they-ate bread and drank beer
'So they ate bread and drank beer.'
(Gr. 129, 1–3; Foulet 1990: 62)
- b. En me bourse grande a il **deniers** a grant planté.
in my purse big has there coins in great plenty
'In my big purse there is money in great plenty.'
(Av. 203–4; Foulet 1990: 63)

We see then that French has lost a class of null indefinite determiners; these were replaced by the indefinite article *un(e)*, the 'partitive article' *du*, *de la*, *des* and, for generic plurals, the plural definite article *les*. In this connection, Déprez (1999: 416) points out that 'an attractive conjecture is that the use of bare *rien* and *personne* in environments from which bare NPs gradually disappeared, survived by ... undergoing incorporation into the obsolete

⁹ Actually *aucun* turned into a D and so retained phi-features. D may be the only position in DP where phi-features are systematically marked in Modern French (Harris 1978: 74–5).

empty indefinite determiners which preceded them'. Roberts and Roussou (2003: 149ff.) develop Déprez's conjecture by supposing, following Longobardi (1994; 2001), that Ds give nominals their referential properties, i.e. their ability to refer to objects or sets of objects in the world. Once the OF null indefinite Ds were lost (which was presumably due to the extension of the use of the indefinite and 'partitive' articles; see Foulet (1990: 54ff.) on these developments), DPs with null Ds could no longer be referential. Words such as *rien*, *personne*, and *point*, as well as a small number of others including *chose* ('thing') and *âme* ('soul') (Foulet 1990: 275ff.), which for a time were also negative expressions, could remain in such DPs, but had to be interpreted as non-referential quantifiers occupying Num. The fact that Nouns like *rien* and *personne* denoted 'generic' entities ('thing', 'person') clearly helped in their reanalysis as quantifiers; in this respect *point* is rather different, being originally a Noun denoting a 'minimal quantity' (a 'minimizer' in the terminology of Bolinger (1972)), a point I return to directly.

This accounts for how these words ceased to be nouns, but it does not account for how they became negative (i.e. took on an interpretable Negative feature in terms of what was proposed in §1.4.2). Roberts and Roussou suggest that this change is bound up with the loss of null indefinite Ds in French, as mentioned above, along with the development of a null *negative* D in examples like (27) (see Kayne (1984: 48ff.) for an analysis of the DP bracketed in (27) as containing a null negative determiner):

- (27) Jean n'a pas mangé [_{DP} e de pommes]
 John neg-has not eaten of apples
 'John has not eaten (any) apples.'

This is the only case of a null D in Modern French, and it is negative.¹⁰ In OF, this construction did not exist; see the detailed discussion in Foulet

¹⁰ Except in indefinites with the 'partitive article' where the head Noun is premodified by an Adjective:

- (i) a. J'ai acheté **du** pain.
 I've bought of-the bread
 'I've bought some bread.'
 b. J'ai acheté **de** bon pain.
 I've bought of good bread
 'I've bought some good bread.'

Kayne (1984: 79) suggests that this *de* is an article, rather than there being a null determiner or a quantifier.

(1990: 73ff., 264ff.). Instead, a singular negative indefinite typically lacked an overt article altogether:

- (28) a. je ne nourriroie [_{DP} trahitor].
 ‘I would not feed [a] traitor.’
 (Ch. 1223–4; Foulet 1990: 73)
- b. [_{DP} Offrande] hui mais n’i prenderai.
 offering today more not-there I-will-take
 ‘I will take no more offerings today.’
 (F. 570; Foulet 1990: 59)

This construction is slightly different from that in (26) as it features singular count nouns in negative contexts, while in (26) we have mass or plural nouns in non-negative contexts. The article-less DPs in (28) are non-specific indefinites.

Roberts and Roussou follow Foulet (1990: 264ff.) in taking the development of the null negative D, associated with *de*, as being caused by the same reanalysis as that which created clausal *point*. (29a) is an example of *point* in a positive context (albeit an *if*-clause, and as such a context for negative-polarity items – §1.4.1) and (29b) is an example of *point* in a negative context. In both cases, it is followed by partitive *de*:

- (29) a. Ja por rien nel te deïsse
 already for nothing not-it you I-would-say
 se **point de ton bien** i veïsse.
 if bit of your goods there would-see
 ‘I would not tell you if I saw the smallest piece of your goods.’
 (P. 7261–3; Foulet 1990: 268)
- b. cel aweule la qui n’a **point d’argent ne de houce** ausi
 that blind-man there who not-has bit of money norof clothes too
 ‘that blind man who doesn’t have a single bit of money nor clothes’
 (Av. 232–4; Foulet 1990: 266)

In the examples above, the verb is transitive, and *point* can be interpreted as the head of the direct-object DP taking a partitive PP-complement. Thus the relevant part of the structure of (29b) would be as follows:

- (30) V[_{DP} [_D∅] [_{NumP} [_{Num} point] [_{NP} (point) [_{PP} d’argent ...]]]]

In this structure, *point*, like *aucun*, *rien*, and *personne* as discussed above, is reanalysed as merged in Num when the loss of the null indefinite D meant that referential Nouns were no longer legitimate in determinerless DPs.

However, two things distinguish *point* from the other elements which were reanalysed from N to Num. The first is a semantic difference: *point*

lacks the descriptive content susceptible of being reinterpreted as a negative quantifier, i.e. it does not have the ‘generic’ meaning of words like *rien* and *personne*. Instead, it is a minimizer in Bolinger’s sense. Second, *point* was able to be syntactically separate from the following *de*-phrase, as in (31):

- (31) De contredit n’i avra point.
 of opposition not-there will-have bit
 ‘There will not be a bit of opposition.’
 (P., 494 and 3946; Foulet 1990: 265)

In examples like (31) the *de*-phrase satisfies the V2 constraint operative in OF (see §1.3.2).¹¹ The syntactic separability of *point* from the *de*-phrase combines with *point*’s lack of semantic content beyond ‘pure’ negation to create the circumstances for the reanalysis of *point* as a clausal negator, and thus the reanalysis of the DP headed by the null article as negative. This second reanalysis, which affected *point* but not *rien* and *personne*, can be schematized as follows:

- (32) a. ne V [DP [D Ø_{non-specific}] [NumP [Num point] [NP d’argent...]]]>
 b. ne V [Neg point] [VP [DP Ø_{negative} d’argent]]

This is the origin of both the null negative determiner and the clausal negator *point*. The other clausal negators *pas* and dialectal *mie* (from the noun meaning ‘crumb’, another minimal quantity) underwent a similar reanalysis. As Foulet (1990: 269) points out, once expressions like *il n’y a pas d’argent* (‘there is no money’) arise, the development of the negative *de*-phrase is complete, since these are etymologically absurd, i.e. they could not mean ‘there is not a step of money’, although negative *pas* derives from a Noun meaning ‘step’.

So we see how the development of the null negative Determiner is connected with the development of clausal negator *point*. The result of this development, combined with the loss of the null non-specific indefinite article of (28), is that null Ds are always inherently negative. Now, since *rien* and *personne* were the only Nouns able to appear with a null Determiner, they too became inherently negative. In terms of the analysis in § 1.4, they took on an interpretable Negative feature. In this way, all three developments – the development of clausal *point*, the development of the null negative D, and the development of *rien* and *personne* as n-words – are linked together by the loss of N-to-Num movement and the reanalysis in (32).

¹¹ Whether the fronted constituent *de contredit* here is a PP, an NP or a DP is not clear, but not crucial for the point at issue.

So we see that this case of grammaticalization involves reanalysis. The reanalysis was caused by the ambiguous expression of the interpretable Negative feature. In the older structure in (32a) the interpretable Negative feature is associated with *ne*, while in the innovative structure in (32b) an interpretable Negative feature is associated with both *point* and the null D of the object DP. The latter two items form a ‘negative concord’ relation, i.e. an Agree relation of the type discussed for Modern French in §1.4.1. It may be that this reanalysis led to Negative Agree relations in French, and thus changed the value of Parameter D of Chapter 1.¹² So grammaticalization can be seen as parameter change with associated reanalysis. The parameter change takes place when the P-expression is ambiguous and a reanalysis happens. Roberts and Roussou (2003) present a number of cases of the same type affecting the T-, C-, and D-systems; in each instance grammaticalization involves reanalysis triggered by ambiguous P-expression and associated reanalysis.

2.3. Argument structure

2.3.1. Thematic roles and grammatical functions

In this section I turn to changes in **argument structure**, the way in which the participants in the action or state of affairs described by a predicate are realized in the structure of the sentences containing that predicate. An important distinction to be made in this connection is that between semantic (or **thematic**) **roles** such as Agent, Patient, Recipient, etc., and grammatical functions such as subject, direct object, indirect object, etc.

¹² However, it is difficult to be sure of the chronology in this case. We saw in §1.4.2 that *ne ... pas* became the obligatory form of negation in the seventeenth century. *Personne* became an n-word in the seventeenth century, and Brunot and Bruneau (1937), cited in Déprez (1999: 414), point out that the changes in the article system were not complete until that time. We may therefore tentatively continue to date the change as taking place at this period. It is quite likely that *ne* had an optionally interpretable Negative feature for some period, as mentioned in §1.4.2. See however Chapter 3, note 7, for some indication that this chronology may not be fully correct.

Following a standard view in generative grammar, I take thematic roles to be primitives associated with each verb¹³ as a matter of lexical semantics (which is not to say that verbs do not fall into lexical classes; they do, as we shall see), and grammatical functions to be defined in terms of structural positions. Thus, the subject is the DP Specifier of TP and the direct object is the DP complement of V, for example. Although distinct, there is a relation between thematic roles and grammatical functions: for example, Agents are always subjects (in active clauses), although subjects need not always be Agents, as the subjects of stative verbs like *know*, *believe*, and *contain* show. The relation between thematic role and grammatical function is specified lexically for each verb.

As just mentioned, verbs fall into lexical subclasses. These can be defined in terms of the number of thematic roles they have and the way in which they distribute these. A thorough discussion of the verb classes of English can be found in Levin and Rappaport-Hovav (1995); here I will limit my attention to the one or two types which are relevant for the discussion of changes to follow. Traditional grammar recognizes a distinction between transitive and intransitive verbs (see Law (2003: 90) on the origin of this notion); the former are verbs with two thematic roles (for example, *eat*, *like*, *hit*) while the latter are verbs with just one (for example, *laugh*, *cough*, *fall*, *die*). There are also verbs with three thematic roles, such as *give*, *send*, *show*. Some of these verbs can appear in what is often called the ‘double-object’ construction as in *John sent Mary a letter* (see Chapter 1, note 9); we will be looking at one change that has affected this construction in the history of English below.

Recent linguistic theory has established a distinction between two types of intransitives: unergatives such as *laugh* and *sing*, and unaccusatives such as *fall* and *die* (see Perlmutter (1978); Burzio (1986)). In the former, the single argument of the verb is a true subject. (These verbs are usually agentive.) In the latter, the verb’s argument is merged as an object, and moves to the subject position. There is much cross-linguistic evidence for this distinction, although in English the evidence is rather indirect. The clearest indication of unaccusativity in English lies in the availability of a deverbal adjective formed from the verb’s participle: thus we have *a fallen leaf*, meaning ‘a leaf which has fallen,’ but not *a laughed man* (meaning ‘a man who has laughed’); so we see that *fall* is an unaccusative verb and *laugh* is an unergative.

¹³ Actually all lexical categories assign thematic roles, but I restrict attention here to verbs as this is the richest category in terms of thematic structure, and also because the changes we will be looking at concern verbs.

The distinction between thematic roles and grammatical functions can be observed when we compare agentive transitives with so-called ‘psychological’ verbs (henceforth psych verbs), i.e. those which describe a psychological event or state. Consider the following pair of sentences:

- (33) a. John reads the newspaper.
b. John likes the newspaper.

In both of these examples, *John* is the subject and *the newspaper* is the direct object. However, while in (33a) *John* is the Agent of the action described by *read* and *the newspaper* is the Patient of the action, in (33b) *John* has the thematic role of Experiencer, the person of whom the psychological state described by *like* holds, and *the newspaper* is what that state is about, the Theme.¹⁴ Psych verbs, unlike action transitives, can in fact distribute their thematic roles ‘the other way around’, as it were, making the Theme the subject and the Experiencer the object: compare *the newspaper pleases/amuses/annoys/appals John* with (33b). This possibility gives rise to doublets of psych verbs which are very close in meaning but which distribute their thematic roles differently, such as *like/please, fear/frighten*, etc. Many languages have a third psych-verb construction in which the Theme is the subject and the Experiencer is marked like an indirect object. This construction is restricted to one verb in present-day English, *appeal* (as in *the newspaper appeals to John*), but is cross-linguistically common. The (near) loss of this construction is one important change in argument structure that we will look at below.

Argument structure can be manipulated by syntactic operations. The best known and probably most widespread such operation is the passive. In the passive of a transitive verb, the DP which corresponds to the direct object in an active sentence functions as the subject and the DP corresponding to the subject of an active sentence is either absent or appears in a *by*-phrase, as in *the newspaper is read (by John)*. Double-object verbs passivize the first object, which in fact corresponds to the notional indirect object, as the following examples illustrate.¹⁵

- (34) a. John sent Mary a letter.

¹⁴ The terminology associated with thematic roles is notoriously varied. I will attempt to use the most neutral labels possible, and hence use ‘Theme’ here. Pesetsky (1994: 56ff.) argues that there are in fact various thematic roles associated with what I am calling the Theme argument of psych verbs.

¹⁵ The ‘%’ diacritic in front of (34e) indicates that the example is not acceptable in all dialects of English. Most American speakers reject examples of this type. They may be more natural in Northern varieties of British English than in Southern ones.

- b. John sent a letter to Mary.
- c. Mary was sent a letter (by John).
- d. A letter was sent to Mary (by John).
- e. %A letter was sent Mary (by John).

The passive in (34c) is known as the ‘recipient passive’. This construction has changed in the recorded history of English, as we shall see below.

I follow the standard assumption in taking both thematic roles and grammatical functions to be universal. Languages vary somewhat in how grammatical functions are morphosyntactically marked: the Modern English system relies primarily on word order, but many languages have morphological case marking on DP constituents (nouns, articles, and other DP-internal elements such as adnominal adjectives), which plays a major role in marking grammatical function; this was the situation in Latin and, to some degree, in OE. Still other languages may mark grammatical functions by means of agreement, and many languages combine these various methods. The pattern of case/agreement marking in relation to grammatical functions may also vary: in §4.1, we will look at the concept of ergativity (see Box 4.1). We take all of this to involve parametric variation (concerning the Agree and Move relations), retaining the view that grammatical functions are defined in structural terms however they are overtly marked.

2.3.2. *Changes in English psych verbs and recipient passives*

Where there is synchronic variation, there is diachronic change. The ways in which languages mark grammatical functions can change, as indeed they have done in the history of English and in the development from Latin to Modern Romance. Here I want to focus on two changes involving the marking of grammatical functions in the history of English. I will suggest that one of these changes, at least, is a parametric change associated with reanalysis. The other change may be of a different nature, being a change affecting the lexical properties of verbs, although the parametric change is also relevant to it. The first change concerns recipient passives and the second concerns psych constructions. Both have been much discussed in the recent literature on diachronic syntax: see Allen (1986; 1995); Anderson (1986); Denison (1990; 1993: 103ff.); Fischer and van der Leek (1983); Lightfoot (1979; 1991: 128ff.; 1999: 125ff.); the main traditional studies are van der Gaaf (1904) and Jespersen (1909–49, III). Much of the

following discussion is based on Allen (1995), which is the most thorough and authoritative study of these and other changes.

The first change affects recipient passives. (35) shows that in recipient passives in OE the subject of the passive retained its indirect-object marking, i.e. dative case:

- (35) ac **him** næs getiðod ðære lytlan lisse.
 and him-Dat not-was granted the small favour-Gen
 ‘But he was not granted that small favour.’
 (*ÆCHom* I 23.330.29; Denison 1993: 108)

Here the other argument of ‘grant’, *ðære lytlan lisse*, is in the genitive case; I will return to this point below.

Concerning the second change, one class of psych constructions is illustrated in (36). This is the third type of psych construction described above – that where the Experiencer is marked as an indirect object. (36) shows that this construction was found in OE:

- (36) hu **him** se sige gelicade
 how him-Dat the victory-Nom pleased
 ‘how the victory had pleased him’
 (*Or* 84.32; Denison 1993: 72)

Here the Experiencer, *him*, is in the dative case, the typical marking of an indirect object in OE (the accusative 3sg masculine pronoun was *hine* in OE). The Theme argument is in the nominative case, although it was not the subject; see Allen (1995) for extensive discussion of this point. The verb, translated as ‘please’ here, is *lician*, the ancestor of NE *like*. We observe that this verb has undergone a redistribution of its thematic roles (although not really a change in meaning, since the core meaning has involved causing pleasure all along), presumably associated with the loss of the construction in (35).

A further point that is relevant here is that certain two-argument verbs required their object to have some case other than the accusative. When passivized, the case of the active object is retained on the passive subject. This can be seen in the following example with the passive of ‘help’, a verb whose object is required to be dative:

- (37) Ac **ðæm** mæg beon suiðe hraðe geholpen from his lareowe.
 and that-Dat may be very quickly helped by his teacher
 ‘But that may be remedied very quickly by his teacher.’
 (*CP* 33.225.22; Fischer *et al.* 2000: 42)

The three constructions shown in (35)–(37) were all lost during the ME period. Since they all feature dative case, and English lost its morphological case system during the same period, it is tempting to relate these developments, as many authors have done. Aside from their intrinsic interest, this lends some importance to the discussion of these changes, as they may reflect some of the syntactic consequences of the loss of morphological case marking.

The NE counterparts to the OE constructions in (35)–(37) are as follows:

- (38) a. How he liked the victory.
b. But he was not granted that small favour.
c. But that may be helped very quickly by his teacher.

Allen (1995) argues that the change affecting psych verbs was distinct from that affecting the passive constructions. Regarding the psych verbs, she says ‘the loss of case distinctions did not make the “impersonal” constructions impossible, but contributed to the decline in frequency of these constructions which ultimately resulted in changes to the grammar which made them ungrammatical’ (12). Regarding the passive constructions in (36) and (37), on the other hand, she states that these changes ‘support the generative view that a syntactic change can be an essentially sudden reanalysis or change in parameter-settings which take place as a by-product of another change which removes . . . the evidence available to language learners for the old analysis’ (446). This contrasts with the ‘lexically-implemented’ change involving the loss of the psych construction (and the associated changes in the relation between thematic roles and grammatical functions in some verbs such as *like*).

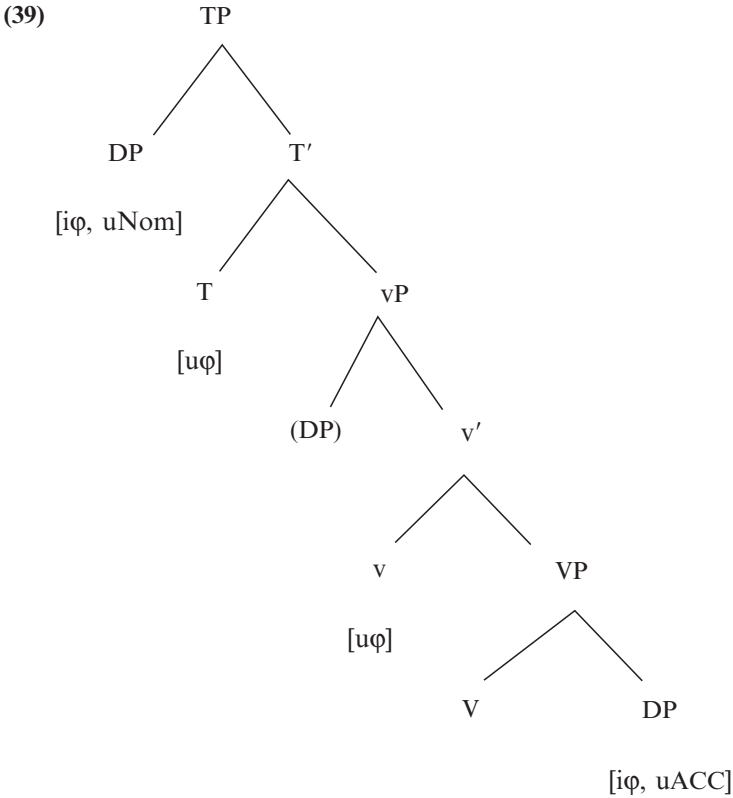
The changes in the psych verbs were thus changes in lexical entries of individual verbs, which diffused through the lexicon over a considerable period. Allen (1995: 221ff.) argues that the beginnings of this change may be discerned in the optional assignment of ‘lexical case’ (for the moment this can be taken to mean dative case, see below) to the Experiencer arguments of certain verbs in OE (for example, *sceamian* ‘to be ashamed’; Allen gives the full range of data in her Table 4.14, 137) and says that ‘while the loss of morphological case distinctions may well have exacerbated the tendency to treat Experiencers as nominative subjects (at least as an option), it did not create it’ (287). The change was completed only by approximately 1500, in that the sixteenth-century examples of this

construction appear to be fixed expressions, the best known of which is *methinks*.

We are thus led to observe a non-parametric kind of change involving changes in lexical properties of verbs diffusing through the lexicon over a long period. I will defer a fuller discussion of implications of this point to §4.1. On the other hand, as Allen says, the changes to the passive constructions are candidates for treatment as parametric change. I will suggest that there was a parameter change and associated reanalysis involving the interpretable and uninterpretable case features associated with the DPs bearing thematic roles (the arguments) in these constructions. This parameter change was associated with a reanalysis caused by the loss of overt morphological case distinctions.

OE had a morphological case system in which nominative, accusative, dative, and genitive cases were distinguished. There was already some syncretism among case forms in OE, and Allen (1995: 158ff.) shows in detail that the system had broken down in all the ME dialects except Kentish by the end of the thirteenth century at the latest (see her Table 10.1, 441). Allen argues that this change directly caused the loss of indirect passives in the early thirteenth century: '[t]he indirect passives [i.e. the construction in (35) – IGR], disappear just at the time when this morphological distinction disappeared in most of the country and follows straightforwardly from the fact that there was no longer any evidence available to language-learners for two types of objects of monotransitive verbs' (446).

We can understand this change as a reanalysis and associated parameter change, but in order to see this, certain assumptions about case and arguments must be introduced. Generative theory postulates the existence of abstract Case (written with a capital 'C' to distinguish it from morphological case). Nominative and Accusative Case can be thought of as uninterpretable features associated with argument DPs which must be deleted under an Agree relation with features of a relevant head, subject to the usual conditions on Agree (see §1.4). Nominative Case is deleted under Agree with the φ -features of finite T, giving rise to subject-verb agreement in finite clauses. Accusative Case is similarly deleted under Agree with v, a verbal functional category which takes the lexical VP as its structural complement. Both the Probe and the Goal bear uninterpretable features in these instances. The φ and Case features in a simple transitive sentence (for example, *John loves Mary*) are thus as in (39):



Here we have the Agree relations described in (40):

- (40) a. v 's $[u\phi]$ features Agree with those of the object, resulting in deletion of v 's features and of DP's ACC feature;
 b. T's $[u\phi]$ features Agree with those of the subject, resulting in deletion of T's ϕ -features and DP's NOM feature.

In addition, the subject DP raises to SpecTP. We will return to the mechanism which causes Move to happen in §2.5. For now, it is important to note that the subject is merged in SpecvP. From now on, I will adopt this version of the idea that the subject is merged in a predicate-internal position, rather than the idea that the subject is merged in SpecVP, which was discussed in §1.3.1. It is important to see that the Case/Agree relations here are purely structural, in that they are completely blind to the lexical properties of the verb.

Burzio (1986: 178) puts forward an important generalization regarding the nature of transitive clauses: Accusative Case is present if and only if the

verb assigns a subject thematic role. This idea has become known as Burzio's generalization. Assuming a category such as *v* makes it possible to locate the two properties related by Burzio's generalization on a single head. (I will refer to these as the 'Burzio properties' henceforth.) Let us suppose that *v* assigns the subject thematic role and, as just outlined, is responsible for deleting the uninterpretable Accusative feature of the direct object. Following Chomsky (2001:6), when *v* has this double property we write it as *v**.

In unaccusative clauses, *v* is either defective or absent; either way, it lacks the 'Burzio properties' of Agreeing for Accusative Case and having a thematic subject merged in its specifier. I will assume for simplicity that *v* is simply absent in these cases. (This will be revised in §4.1.3.) Under these conditions, the object can be marked as Nominative and raises to the subject position:

(41) [_{TP} John+NOM T[_{uφ}] [_{VP} arrived (John+NOM)]]

In passives, *v* also lacks the Burzio properties. This is what causes the object to be able to appear in the subject position, and the subject to be demoted:

(42) [_{TP} John+NOM [_{T[_{uφ}]} was] [_{VP} *v* [_{VP} arrested (John+NOM)]]]

Presumably the passive morphology on the verbal participle plays a crucial role in determining the 'defective' nature of *v* here – see Baker, Johnson, and Roberts (1989) for more on this point; it is possible that the passive participle raises to *v*, although I will not indicate that here. (Collins (2005) develops a somewhat different analysis of passives.)

Now, let us suppose that languages with richer morphological case marking than NE, such as OE, make syntactic distinctions among the features which license the verb's arguments in addition to the simple Accusative vs. Nominative Case of NE.¹⁶ As the OE evidence we have seen clearly shows, we need to allow for Dative arguments of *V*, i.e. an abstract Dative Case which corresponds to morphological dative case. (We also saw an example of a Genitive argument in (35); the same considerations apply here.) 'Inherent' Cases of this type are known to be directly associated with thematic roles in a way in which structural Nominative and

¹⁶ NE also has a Genitive Case, operative inside DPs, but I will leave that aside here and restrict the discussion to Case at the clausal level.

Accusative are not; the presence of these cases is determined by the lexical properties of the verb. Because of this, we can think of features such as abstract Dative Case (or DAT) as interpretable Case features, or in other words as the morphological realization of a thematic role. As such, these Case features need not and cannot be deleted under Agree.

Now we are in a position to see what the parametric change was in thirteenth-century English. In OE and EME, as long as the morphological case distinctions were manifest, *v* Agreed with abstract Accusative Case, corresponding to morphological accusative case only: other arguments of *V* bore interpretable Case features which as such required no Agree. After the loss of morphological case distinctions, *v*'s feature makeup changed in such a way that it Agreed with any and all non-subject arguments, i.e. its φ -features valued the Case feature of any available DP. This parametric change was associated with the reanalysis shown in (43):¹⁷

- (43) [_{CP} Him+DAT [_{TP}[_T was] [_{VP} v [_{VP} helped (him+DAT)]]]] >
 [_{TP} He+NOM [_T[_{u φ} was] [_{VP} v [_{VP} helped (John+NOM)]]]]

In the new structure in (43), the passive *v* has a different status, in that its ability to value Case on the object has been switched off, while in the old structure *v* had no such property to be switched off; this was a variety of impersonal passive in the sense that no internal argument needed to have its Case feature deleted as the only feature available was the interpretable DAT feature. The head *v* here is thus intransitive, as it is in unergative intransitives, or when the only complement of *V* is a PP or a CP. I attribute the cause of reanalysis to the loss of morphological case distinctions, exactly as stated by Allen (1995: 446): it 'follows straightforwardly from the fact that there was no longer any evidence available to language-learners for two types of objects of monotransitive verbs' (446). Once *him+DAT* could no longer be distinguished from *he+NOM*, the latter option was chosen and, by implication, *v* was taken to be 'personally passive' in this kind of example. This also led to subject agreement, since *T* has

¹⁷ Dative Experiencers could be subjects in OE and ME, as in Modern Icelandic (Sigurðsson 1989); Allen (1995: 50ff.) argues extensively that some preverbal datives in OE were subjects. On the other hand, Allen (1995: 143) points out that 'although PDEs [preposed dative Experiencers – IGR] behaved like subjects, preposed dative Recipients in passive ditransitive constructions did not'. Accordingly, I treat the root category as CP here, with the Dative argument a topic in SpecCP (the internal argument of *help* is a Recipient, not an Experiencer, but this does not affect the syntax). See also Eythórsson and Barðdal (2005: 842–3).

taken on subject φ -features as a consequence of this change. In the older structure, T was ‘impersonal’, i.e. its φ -features bore the default 3sg values.

This change affected v 's feature makeup, and as a parametric change had a number of consequences. The first was the loss of indirect passives. As Allen shows, this construction was lost in the early thirteenth century. Since v took on uninterpretable φ -features in all (transitive) clauses, a further consequence was the loss of all non-subject arguments which formerly bore inherent Case (either Genitive or Dative), i.e. all arguments in the c-command (and therefore the Agree) domain of v . Allen points out in connection with genitive-marked object arguments that ‘we can say that in those dialects in which there is clear evidence of a dative/accusative distinction, genitive objects are still found, while in those dialects in which this distinction was lost, genitive objects are not found’ (217). (She does go on to point out that genitive objects were in any case being replaced by accusative or prepositional objects from late OE onwards). Third, dative-marked subject arguments can survive; Allen (1995: 221ff.) argues that exactly this happened in the case of many psych verbs. A fourth and related point is that this development contributed to the loss or lexical reorganization of psych verbs. As we have seen, the changes in psych verbs were gradual, but the parametric change in v 's properties ruled out a formerly available possibility for these verbs and so played a role in furthering the ongoing lexical changes.

Finally, the reanalysis in (43) affected recipient passives. However, Allen shows that the modern construction is not reliably attested before 1375, over a century after the parametric change just discussed. She comments ‘the historical record does not support the notion of a *replacement* of the dative-fronted passive by the recipient passive ... Rather, the dative-fronted passive seems to have died out from the texts some time before the recipient passive was introduced’ (447). There is, in fact, a period in the fourteenth century when neither construction is found. The parameter change we have proposed will account for the disappearance of the old, dative-fronted construction, since this contained a Dative argument in the c-command domain of v in the thirteenth century. But in itself it predicts nothing about the introduction of the modern-style recipient passive. In fact, the NE recipient passive requires two occurrences of v^* , one associated with the subject thematic role and uninterpretable φ -features, and the other with the indirect-object thematic role and uninterpretable φ -features, as shown in (44) (on the motivation for this type of analysis of double-object constructions, see Larson (1988)):

- (44) [TP John [_{v*P} (John) v* [u ϕ] [_{VP} sent [_{v*P} Mary+ACC v*[u ϕ] [_{VP} V a



Agree

letter+ACC]]]]



Agree

Here the upper v*'s [u ϕ] features Agree with those of *Mary* and the lower v*'s Agree with the ϕ -features of *a letter*. In modern recipient passives, only the upper v* is 'switched off' in the sense mentioned above:

- (45) [TP John [_T was] [_{VP} v [_{VP} given [_{v*P} t(John) v* [u ϕ] [_{VP} V a book+ACC]]]]]



Agree

The lag in the appearance of the modern-style recipient passives in the fourteenth century identified by Allen may be attributable to some further condition required for the innovation of this structure, perhaps connected to the fact that it features two occurrences of *v* (cf. Allen's (1995: 448) comment that this construction appears at the time when the relative order of the indirect and direct objects becomes fixed in this construction).

Finally, let us briefly reconsider the psych verbs. The reanalysis affecting a verb such as *like* must have had the form in (46):

- (46) [_{CP/TP} Him+DAT ... T[u ϕ] [_{VP} (him) like pears+NOM]] >
 [_{TP} He+NOM T[u ϕ] [_{v*P} (he+NOM) v*[u ϕ] [_{VP} likes pears+ACC]]]

Assuming that the OE psych construction was a kind of unaccusative (see Belletti and Rizzi (1988) on this), there would be no *vP* in the structure. This would have permitted *T* to Agree for *NOM* with the direct object, as shown in the first line of (46). (Bejar (2002: 314, 317) proposes a similar analysis for these constructions.) The reanalysis involves the introduction of *v*P* into the structure, with its normal properties of being associated with a subject argument in its Specifier and Agreeing with the *VP*-internal object. (Again, Bejar (2002: 323, 325) proposes the same thing.) In line with Allen's (1995) conclusions, as reported above, I take it that this reanalysis was facilitated, but not caused, by the parameter change affecting the feature-content of *v*.

In conclusion, in this section we have looked at well-known examples of changes affecting both the realization of argument structure of psych verbs and changes in the functioning of the major grammatical-function changing operation, the passive. Largely thanks to Allen (1995), these changes are empirically quite well-documented. We have seen how a parameter change and associated reanalysis can account for many aspects of these changes, in line with the general approach being advocated in this chapter. As it affected the history of English, this change is usually thought of as an example of how the loss of morphological case marking may affect syntax and the lexicon. In the parametric analysis sketched above, this idea is directly reflected, in that v^* is associated with φ -features in systems where the complements of the verb do not show morphological case distinctions. Where morphological case distinctions are found, v acts essentially as in intransitives, lacking these φ -features since the internal argument has an interpretable Case feature which does not require valuing.

The most important general conclusion from the above discussion is that we have observed the interaction of two kinds of change: a parameter change (the change in the feature content of v^*) and a series of changes affecting the lexical entries of psych verbs which diffused through the lexicon over a long period. Although the parametric change influenced the lexical change, the two changes are in principle independent and operate in rather different ways.

2.4. Changes in complementation

In this section we are once again concerned with the nature of the arguments bearing thematic roles determined by verbs. However, the focus here is not on changes in how thematic roles are mapped onto the grammatical functions or on changes in grammatical-function changing operations such as passives, but rather on how the same argument in the same grammatical function may change status. Moreover, I will concentrate on arguments that express a proposition of some kind, and which are therefore typically realized as clausal constituents (mostly but not exclusively as CPs). So the main focus will be on how the propositional arguments associated with verbs like ‘order’, ‘desire’, ‘say’, etc. may change their syntactic properties, without changing either their thematic role (roughly Theme, in each of these cases) or their grammatical function (structurally the complement of the verb in each of these cases).

The particular example of changes in complementation I will look at here are those which distinguish the Romance languages as a whole from Latin. Given the range of languages and constructions to be discussed, the treatment will of necessity be rather coarse-grained. Nevertheless it is possible to observe some interesting diachronic processes at work, and once again we encounter reanalysis and, arguably, parametric change. The general conclusion will be that these notions are relevant to the diachrony of (clausal) complementation, as they are to the other diachronic processes discussed in this chapter.

Vincent (1988: 65–7) summarizes the clausal complementation system of Latin, presenting five main types of complement, as follows:

- (47) a. *ut/ne* + subjunctive (verbs of ordering, desiring, warning, requesting, urging, fearing, etc.):¹⁸
 Ubii Caesarem orant ut sibi parcat.
 ubii-Nom Caesar-Acc beg-3pl UT selves-Dat spare
 ‘The Ubii beg Caesar to spare them.’
- b. (Bare) infinitive (‘want’, ‘prefer’, ‘dare’, ‘try’, ‘begin’, etc.):
 Volo vincere.
 want-1sg to-win
 ‘I want to win.’
- c. Accusative + infinitive (‘verbs of saying, thinking, hoping, perceiving’ (67)):
 Dicit te errare.
 says-3sg you(sg)-Acc to-go-wrong
 ‘He says you are going wrong.’
- d. *Quod* (or *quia*) + indicative (‘verbs of emotion where in a loose sense the complement can be said to express the cause or origin of the emotion’ (67)):
 Dolet mihi quod tu nunc stomacharis.
 pains-3sg me-Dat QUOD you(sg)-Nom now are-angry-2sg
 ‘It pains me that/because you are angry now.’
- e. Indirect question (‘any verb with the appropriate meaning’ (67), marked by an initial *wh*-expression in the subordinate clause with the verb in the subjunctive):

¹⁸ There is a further class of complements in *ut*, following mostly impersonal verbs expressing existence, non-existence or simple events:

- (i) Accidit ut esset luna plena.
 happened UT be-imperfect.subjunc-3sg moon full
 ‘There happened to be a full moon.’

(*B.G.* 4, 29, I; Ernout and Thomas 1993: 304)

This *ut* is negated with *ut non* rather than *ne*. I will leave it aside in what follows; see Ernout and Thomas (1993: 303ff.)

Ab homine quaesivi quis esset.
 from man-Abl asked-1sg who be-3sg-subjunc-imperf
 'I asked the man who he was.'
 (Ernout and Thomas 1993: 313)

In most of the Romance languages, in particular French and (Standard) Italian, this system has changed quite drastically. This can be seen if we translate the above examples into French:

- (48) a. Les ubii supplient César de les épargner.
 the Ubii beg-3pl Caesar DE them spare
 'The Ubii beg Caesar to spare them.'
- b. Je veux gagner.
 I want to-win
 'I want to win.'
- c. Il/elle dit que tu te trompes.
 s/he says-3sg that you(sg)-Nom you(sg)-Acc mistake
 'He says you are going wrong.'
- d. Ça me fait de la peine (parce) que tu es fâché maintenant.
 it me-Dat makes of the pain (because) that you(sg) are angry now
 'It pains me that/because you are angry now.'
- e. J'ai demandé à l'homme qui il était.
 I have asked to the man who he was-imperf
 'I asked the man who he was.'

The changes can be summarized as follows (see Vincent (1988: 68)):¹⁹

- a. loss of *ut/ne* + subjunctive;
- b. restriction in distribution of the bare infinitive;
- c. loss of accusative + infinitive;
- d. the spread of *quod*-clauses into former (c) environments;
- e. no change (except that mood of lower clause may now be indicative) in *wh*-clauses.

The different changes illustrate a variety of patterns of loss, restriction, spread and, in the case of (e), (near) stability. Let us look at the changes more closely and see whether we can see any more general patterns.

a. In the Modern Romance languages, the Latin *ut* + subjunctive construction has completely disappeared and has been replaced by 'prepositional infinitives', i.e. infinitival clauses introduced by a complementizer

¹⁹ Vincent adds the development of the causative construction from *facere* ('do/make') + infinitive. I will leave this construction aside here, since it arguably also involves changes in grammatical functions.

derived from a preposition, *a/à* or *di/de*. Prepositions frequently grammaticalize as complementizers; this kind of development is discussed in Haspelmath (1989), Hopper and Traugott (2003: 188–90), Roberts and Roussou (2003: 97ff.) and the references given there. It seems quite reasonable to treat *ut* and *ne* as complementizers in Latin, members of category C heading the CP complement of the relevant classes of verbs. Modern Romance *a/à* and *di/de* clauses are usually treated as CPs (see for example, M. Jones (1996: 59), but see Kayne (2000: 282ff.) for a very different view), and we can take *a/à* and *di/de* to be complementizers. In that case, the change that has taken place here seems to involve one type of CP (non-finite, introduced by a grammaticalized preposition) replacing another (finite, introduced by a particle). It is important to bear in mind that ‘replacement’ does not imply ‘reanalysis’ here; it is not clear whether or when the infinitival constructions replaced the *ut* ones, although there is some evidence for complementizer *a* in Vulgar Latin (Gamillscheg (1957: 462), cited in Hopper and Traugott (2003: 189)).

b. The change affecting bare infinitives also involves prepositional infinitives. The latter have replaced bare infinitives in various contexts, notably **object control**, i.e. cases where the reading of the understood subject of the infinitive is determined by the object of the main verb. In Latin, this construction could involve bare infinitives, but in Modern French and Italian a preposition is always required in these cases:

- (49) a. Ab opere ... legatos discedere vetuerat.
 from work-Abl legates move-away had-forbidden-3sg
 ‘He had forbidden the legates to move away from the work.’
 (Caesar, *B.G.* 2, 20, 3; Ernout and Thomas 1993: 329)
- b. Il avait défendu aux légats de s’éloigner des travaux.
 he had forbidden to-the legates DE selves-distance from-the works
 ‘He had forbidden the legates to move away from the work.’
 (Ernout and Thomas’ (1993: 329) translation of (47a))

Similarly, a number of the **subject-control** verbs (i.e. verbs with infinitival complements whose implicit subject is understood as corresponding to the main-clause subject) listed by Ernout and Thomas (1993: 328) require a prepositional infinitive in Modern Romance, for example, *studeo* (‘be eager’), *postulo* (‘claim’). In fact, the verbs which take a bare infinitival complement in Modern Romance fall into rather restricted classes: ‘semi-auxiliary’ verbs (for example, *vouloir* (‘want’), *pouvoir* (‘can’), the causatives *faire* (‘do/make’) and *laisser* (‘let’), perception verbs such as *voir*

(‘see’), *entendre* (‘hear’), and some impersonals such as *falloir* (‘be necessary’). In all of these cases, it is likely that the complement clause is somehow reduced, i.e. not a CP but perhaps a TP or vP. (See Wurmbard (2001) for an exploration of a variant of this idea.)

One important class of verbs with a propositional bare-infinitive complement in Modern French is the so-called ‘cognitive verbs’, which express ‘belief or the communication of belief’ (M. Jones 1996: 414). These in fact correspond to the Latin accusative + infinitive construction, where the subject of the infinitive was overt and marked accusative, unlike in French, where the subject is understood and ‘controlled’ by the main-clause subject:

- (50) a. Te abisse hodie hinc negas?
 you(sg)-Acc go-away-past-infin today here deny-2sg-pres
 ‘Do you deny that you left here today?’
 (Vincent 1988: 70)
- b. Est-ce que tu nies être parti d’ici aujourd’hui?
 is it that you deny to-be left from-here today
 ‘Do you deny that you left here today?’

It seems that this construction derives from the earlier accusative + infinitive construction, which in these cases was replaced by a bare-infinitive construction with subject control.

c. The Latin accusative + infinitive construction has disappeared with propositional complements of the type illustrated in (47c) in Romance.²⁰ This construction appears to resemble the English construction in (51), although in Latin it is found in the complement of a wider range of verbs:

- (51) I believe him to be mistaken.

In the construction in (51) the Accusative Case of the subject of the infinitival clause depends on the verb (or, more precisely, v^* – see the previous section) of the main clause. The clearest way to see this is by passivizing the main verb, in which case the subject of the infinitive becomes the subject of the main clause:

- (52) He is believed to be mistaken.

This is what we expect if the Case of the infinitival subject depends on the main-clause v^* . Under passivization, as we saw in the previous section, v^*

²⁰ It survives with perception verbs and causative *laisser*. Here again, though, it is not clear that the complement is a full CP. In any case, it arguably denotes an event rather than a proposition, as argued by Guasti (1991: 36ff, 120ff.).

is rendered ‘defective’ and as such unable to Agree with φ -features on any category in its c-command domain. The subject of the infinitive can thus not be Accusative; instead, it Agrees for Nominative with the main-clause finite T (and moves to the main-clause subject position).

In Latin, it seems we can find the same pattern as in English (51) and (52). In particular, we find examples where the main verb is passivized and the subject of the infinitive appears in the nominative case:

- (53) a. Galli dicuntur in Italiam transisse.
 Gauls-Nom say-passive-3pl in Italy-Acc to-have-crossed
 ‘The Gauls are said to have crossed into Italy.’
 (Ernout and Thomas 1993: 327)
- b. Traditur Homerus caecus fuisse.
 report-passive-3sg Homer-Nom blind-Nom to-have-been
 ‘Homer is reported to have been blind.’
 (Vincent 1988: 67)

We can thus analyse these examples along the same lines as the English one in (52). This implies that in the active accusative + infinitive constructions the Accusative Case of the subject of the infinitive Agrees with v^* of the main clause. (This construction is usually analysed as a TP rather than a CP complement in English (for example, in Chomsky (2001: 8)); below I will suggest that both the Latin and the English constructions are CPs, following Kayne (1984)).

However, the possibility illustrated in (53) was restricted to a subclass of the verbs of saying. (Woodcock (1959: 22) gives an indication of which authors used which verbs in this construction.) The apparently more productive option features the subject of the infinitive in the Latin construction in the accusative independently of the main clause. This evidence for this is summarized in Bolkestein (1979). First, alongside examples like (53), we find examples where the main verb is passive and yet the subject of the complement infinitive is nevertheless accusative:

- (54) Dicitur Gallos in Italiam transisse.
 say-passive-3sg Gauls-Acc in Italy-Acc to-have-crossed
 ‘It is said that the Gauls have crossed into Italy.’
 (Ernout and Thomas 1993: 327)

Second, accusative + infinitive clauses are found as complements to Nouns, which is impossible with the nominal equivalents of English verbs which appear in the accusative + infinitive construction:

- (55) a. *the belief (of) him to be mistaken
 b. nuntius oppidum teneri
 message town-Acc to-be-held
 ‘the message that the town was being held’
 (Bolkestein 1979: 31)

Third, many verbs which appear in the accusative + infinitive construction do not otherwise have an Accusative object. This is true of *dicere* (‘to say’), as in shown in (56):

- (56) a. Dico te venisse.
 I-say you-Acc come-perf-infin
 ‘I say that you have come.’
 b. *Dico te.
 I-say you
 (Bolkestein 1979: 20)

We must therefore allow for some mechanism of Accusative Agreement inside the infinitival clause, since by assumption the passivized *v* of the main clause cannot be responsible for this in cases like (54) and (56) and neither can the noun *nuntius* in (55b). I propose, following Cecchetto and Oniga (2001: 6), linking this to the fact that infinitivals bear morphological marking of tense/aspect and voice in Latin. Thus, alongside the present active infinitive, for example, *facere* (‘to do’), we have the perfect active *fecisse* (‘to have done’), the future active *facturum esse* (‘to be about do’) and the corresponding passive forms *factum esse* ‘to have been done’ (perfect passive), *fieri* ‘to be done’ (present passive) and *factum iri* ‘to be about to be done’ (future passive) (see Harris (1978: 195)). Although many of these forms are periphrastic and imply a rather complex morphosyntactic analysis which I cannot go into here, the coexistence of synthetic forms like *fecisse* and *facere* suggests that Latin infinitives are significantly different from those of Modern Romance, where no such opposition survives. This is further supported, as Cecchetto and Oniga (2001: 6) point out, by the fact that Latin accusative + infinitive clauses allow the full range of infinitival tense-forms:

- (57) a. Dicunt eum laudare eam.
 say-3pl him-Acc praise-infin-Pres her-Acc
 ‘They say that he praises her.’
 b. Dicunt eum laudavisse eam.
 say-3pl him-Acc praise-infin-perfect her-Acc
 ‘They say that he praised her.’

- c. Dicunt eum laudaturum esse eam.
 say-3pl him-Acc praise-infin-fut to-be her-Acc
 ‘They say that he will praise her.’

Let us suppose that the tensed nature of Latin infinitives implies the presence of a functional head – presumably T – with the capacity to Agree with an Accusative subject. This can explain the data in (54)–(56). (Cecchetto and Oniga (2001: 26) make the same connection between tensed forms of infinitives and the possibility of Accusative subjects of infinitives, but in a technically more indirect way.)

The analysis of (54) just sketched provides a way of understanding why this variant of the accusative + infinitive construction does not survive in Romance. All other things being equal, we predict that it died out with the tense/aspect marking of infinitives, which appears to have died out in Vulgar Latin (Harris 1978: 195). I will return to the question of why the English-style option for the accusative + infinitive was lost. What is clear is that the accusative + infinitive was replaced by clauses introduced by *quod*, the final change to be considered. (Recall that there has been essentially no change in the nature of indirect questions; see (47e) and (48e) above.)

d. The commonest pattern of clausal complementation in Modern Romance involves a finite clause introduced by *que/che*, which derives from Latin *quod*, the nominative/accusative neuter form of the relative pronoun (or perhaps partly from the masculine accusative *quid*; Harris (1978: 228)). Since *que/che* clauses commonly appear as the complements of verbs of saying and believing in Modern Romance, they have clearly taken over much of the distribution of the Latin accusative + infinitive construction. *Quod*-clauses were originally found in various non-complement positions: for example as subjects or adverbials, as in (58):

- (58) a. Multum ei detraxit ... quod alienae erat civitatis.
 much him-Dat detracted ... QUOD foreign-Gen was-3sg city-Gen
 ‘The fact that he was from a foreign city detracted from him a great deal.’
 (Nep. 18, 1, 2; Ernout and Thomas 1993: 295)
- b. Adsunt propterea quod officium sequuntur.
 are.present-3pl on.that.account QUOD duty-ACC follow-3pl
 ‘They are present because they follow duty.’
 (Cicero; Kennedy 1962: 183)

Quod-clauses also followed adverbials such as *nisi* (‘unless’), *praeterquam* (‘except’), etc., and appeared to require a **factive** meaning, in that the truth of the proposition expressed by the complement clause was **presupposed** (see

Ernout and Thomas (1993: 295ff.) and note the factive interpretations of (47d) and (58); the factive interpretation is clearest where the verb is indicative). Still according to Ernout and Thomas (1993: 296), *quod*-clauses appear as complements to ‘metalinguistic’ verbs like *addere* (‘add’), *praeterire* (‘elude’), *mittere* (‘omit’), and with impersonal eventive verbs, usually accompanied by an adverb, or *facere* (‘do/make’) accompanied by an adverb:

- (59) accidit perincommode quod eum nusquam vidisti.
 happened-3sg unfortunately QUOD him nowhere saw-2sg
 ‘It is unfortunate that you didn’t see him anywhere.’²¹
 (Cicero, *At.* 1, 17, 2; Ernout and Thomas 1993: 296)

Here too the factive interpretation is clear.

Finally, both Woodcock (1959: 23) and Ernout and Thomas (1993: 299) point out that *quod*-clauses first appear with verbs of saying and believing in apposition with a neuter form. (Kühner and Stegmann (1955: 270) give examples from Plautus illustrating this development.)

According to Ernout and Thomas, *quod*-clauses appear as a direct complement to verbs of saying and believing only in Vulgar Latin, in Petronius’ imitations of the speech of the lower classes or freed slaves (‘*affranchis ou de petites gens*’), and the language of translations from Greek (following the *légo óti* (‘I say that’) pattern), especially Christian ones. Woodcock (1959: 23) points out that *quod*-clauses commonly appear instead of the Accusative + infinitive ‘from the second century of our era.’ Similarly, Kühner and Stegmann (1955: 279) say that *quod*-clauses replace accusative and infinitives in Late Latin. According to Ernout and Thomas, the earliest example is (60a); (60b) is from Petronius; (60c) is from the Vulgate:

- (60) a. Legati Carteienses renuntiaverunt quod Pompeium in
 legates-Nom from-Carteia announced-3pl QUOD Pompey-Acc in
 potestate haberent.
 power-Abl had-3pl-subjunc
 ‘The legates of the people of Carteia announced that they had Pompey in
 their power.’
 (*B. Hisp.* 36, 1; Ernout and Thomas 1993: 299)
- b. Scis quod epulum dedi.
 know-2sg QUOD meal-Acc gave-1sg
 ‘You know that I gave a meal.’
 (Petronius 71, 9; Ernout and Thomas 1993: 299)

²¹ Ernout and Thomas’ French translation is ‘il est très malheureux que tu ne l’aies vu nulle part.’

- c. Scimus quia hic est filius noster.
 we-know QUIA this is son our
 ‘We know that this is our son.’
 (Vulgate: *John* 9, 20; Ernout and Thomas 1993: 299)

(*Quia*, ‘because’, was an alternative to *quod* at this stage, at least in ecclesiastical writers – see below).

It seems pretty clear, then, that *quod*-clauses were not true complements to verbs of saying and believing in Classical Latin, although they developed into complement clauses in Vulgar Latin. The factive interpretation of *quod*-clauses, their ability to appear as subjects and the origin of *quod* as a relative pronoun all point to an original status as a nominal. Let us suppose then that *quod*-clauses were DPs in Classical Latin, headed by the D *quod*, which in turn selected a CP (see Kiparsky and Kiparsky (1971); Farkas (1992); Roussou (1991; 1994) on the notion of factives as ‘nominalized clauses’). This structure was reanalysed as a CP with C *quod* in Vulgar Latin, and as such it was able to appear in the complement to CP-taking verbs which in Classical Latin took the accusative + infinitive construction, i.e. verbs of saying and believing. The reanalysis is schematized in (61):

- (61) [DP [D quod] [CP [TP epulum dedi]]] > [CP [C quod] [TP epulum dedi]]

I will return below to the parameter change associated with this reanalysis.

Up to now I have been assuming a generic ‘Romance’ complementation system. However, it is worth pointing out that the system is quite different in a number of Southern Italian dialects. According to Rohlf’s (1969: 190), ‘[F]rom Sicily up to Abruzzo, we see in use a double series of conjunctions’, roughly corresponding to the distinction between Latin accusative + infinitive vs. *ut* clauses in their distribution (see also Manzini and Savoia (2005: 455ff.)). This is illustrated for some of these varieties in (62):

- (62) a. Sicilian: pensu **ca** vèni vògghiu **chi** mmanciassi
 b. N. Calabrian: criju **ca** vèni vuogliu **chi** mmangia
 c. Salentino: crisciu **ca** vène ogghiu **cu** mmancia
 ‘I think he’ll come.’ ‘I want that he eat.’

The *ca* complementizer, corresponding to Latin accusative + infinitive, derives from Latin *quia*, which, as just mentioned, was an alternative to *quod* in Classical and Vulgar Latin. However, it appears that *quia* was more common than *quod* in the relevant contexts in older Latin (Ernout and Thomas 1993: 298), and so it is possible that this different system arose in

the area where Latin had been spoken longer. In this area, *quia*-clauses, reanalysed as in (61), took over from Latin accusative + infinitive, and *quod*-clauses, reanalysed in the same way, replaced *ut*-clauses. A consequence of this is that prepositional infinitives are rarer in these varieties than in ‘Standard’ Romance. In roughly this area, infinitives are highly restricted, occurring only in complements to ‘semi-auxiliary’ verbs like *volere* (‘want’), etc. (see Ledgeway (2000: 70ff.)).²²

Returning to the mainstream Romance system, let us recapitulate the changes we have discussed:

- a. loss of *ut/ne* + subjunctive, replaced by prepositional infinitive;
- b. restriction in distribution of the bare infinitive (except with ‘cognitive verbs’);
- c. loss of accusative + infinitive in propositional complements;
- d. the spread of *quod*-clauses into former (c) environments.

All these changes affected the realization of CPs. We can summarize the situation further by saying that two new complementizers emerged in Vulgar Latin: *quod* (through the reanalysis in (61)) and the prepositional complementizers *a* and *de* (through reanalysis of PPs as CPs). The former took over from the accusative + infinitive, and the latter from *ut* and from many instances of bare infinitives. It is important to see that this does not imply that the Classical Latin constructions were reanalysed as the Vulgar Latin ones; the Vulgar Latin constructions arose through reanalyses we have described and simply replaced the Classical Latin constructions.²³

²² There is a further general complementation pattern in Romance, found in Rumanian and in two dialect areas of the extreme south of Italy (Southern Calabria/North-East Sicily and Salento). Here infinitives are almost entirely absent, and specific particles introduce the subjunctive clauses corresponding to Latin *ut*-clauses, as in the Rumanian examples in (i) and (ii):

- (i) Cred că va veni.
I-believe that will come
‘I believe he’ll come.’
- (ii) Voiu să vină.
I-want Prt come-subjunc
‘I want him to come.’

This system is characteristic of the Balkan *Sprachbund*, and is plausibly attributable to the influence of Byzantine Greek (see Calabrese (1993: 73) on Salentino; Ledgeway (1998) on Southern Calabrian/North-East Sicilian).

²³ It is interesting to observe that *ut* probably underwent a reanalysis of a kind similar to that affecting *quod/quia*, in that it earlier functioned as an adverbial introducing a clause. According to Sihler (1995: 399) *ut*, or *uti*, comes from an

But in that case, what caused the Classical Latin constructions to disappear? We can simply assume that *ut* disappeared through phonological attrition. Regarding the accusative + infinitive construction, however, more needs to be said. This is where parameter change becomes relevant. Kayne (1984: 103ff.) observes that English and French differ in two ways as regards infinitive constructions. English allows accusative + infinitive constructions of the kind seen in (51) and (52) (i.e. where Accusative Case on the subject of the infinitive Agrees with v^* in the matrix clause), while French does not. Second, the English prepositional complementizer *for* itself probes Accusative Case on the subject of the complement infinitive:

- (63) a. [For [him to leave]] would be a mistake.
 b. It's nice for the rich [for [the poor to do the work]].

In French, on the other hand, prepositional complementizers cannot be followed by an accusative subject, or indeed any kind of overt subject. Leaving aside a number technicalities (most of which are in any case irrelevant to the version of minimalism we are loosely adopting here), we can formulate the following parameter:

G. does L allow accusative subjects in SpecTP of a non-finite clause?

We can see that Classical Latin and English have a positive value for this parameter, while the Modern Romance languages (quite uniformly, despite all the other differences in their complementation systems) do not.²⁴ A positive value for this parameter allows the English-style accusative + infinitive construction and requires prepositional complementizers introducing infinitives to be followed by overt subjects. Classical Latin in fact lacked this construction, as it did not have prepositional complementizers – these were a Vulgar Latin innovation; Kayne (1984: 117) observes these similarities and differences between Classical Latin and English. A negative value bans the English-style accusative + infinitive construction and overt subjects of infinitives introduced by prepositional complementizers. Technically, the parameter must be stated as an abstract property of C in facilitating or impeding Agree relations.

earlier $*k^wuta$, an indefinite/wh pronoun as the initial labiovelar indicates. (The initial k^w was lost by reanalysis of the negative form $*ne+cut(e)i$ as $nec+uti$). The original meaning was ‘where, so that, as’. In being replaced by reanalysed *a/de* we observe a further case of a ‘cycle’ of grammaticalization (see §2.2).

²⁴ Recall that I am assuming that the complements to causative and perception verbs in Modern Romance are not CPs – see Chapter 1, note 15.

The value of Parameter G changed between Classical and Vulgar Latin.²⁵ This was caused by the reanalysis of *a/de* as C-elements and the loss of T's ability to Agree for Accusative inside infinitival complements, which we suggested above was related to the loss of tense/aspect distinctions in infinitives. The consequences of this parameter change were: the complete loss of the accusative + infinitive propositional complements and the associated reanalysis of factive *quod*-clauses as CPs, along with the reanalysis of accusative + infinitive clauses with a subject coreferent to the main clause as bare infinitives. So here we see the role of reanalysis and associated parameter change in changes in complementation.

The development of the prepositional infinitives was, however, a separate change (which played a causal role in relation to the one just described). This change seems to be connected to a separate parametric change concerning the nature of C. After this change, overt elements in C mark finiteness (*a/de* mark a non-finite CP and, after the change in the status of *quod*, this element marks a finite CP). In Classical Latin, however, overt complementizers marked the mood of the clause: *ut/ne* marked the clause as subjunctive, and *wh*-complementizers ((47e) above), prescriptively at least, were always followed by a subjunctive. Ernout and Thomas (1993: 313–5) point out that indicative indirect questions are found in Plautine Latin and in Vulgar Latin.²⁶ The fact that indirect questions clearly show up in the indicative in Vulgar Latin can be considered a further consequence of this parametric change. One cause of the reanalysis of prepositions as complementizers, as often pointed out (see for example Harris (1978: 198)) was the growing use of prepositions to mark case relations as the morphological case system began to suffer phonological erosion. This was particularly relevant as gerunds and supines, both clausal constructions in Classical Latin, required case. Hence, with the erosion of case marking and its replacement with prepositional constructions, prepositions began to be used with certain kinds of non-finite clauses. This may have facilitated the reanalysis of certain prepositions as complementizers.

So we see how two parameters, both concerning the nature of the category C, may have changed between Classical and Vulgar Latin in

²⁵ It changed in the opposite direction in late ME, perhaps as a consequence of OV > VO word-order change. Fischer *et al.* (2000: 214ff.) provide a very interesting discussion of these developments.

²⁶ Recall that we have suggested that *quod* was not a complementizer in Classical Latin.

such a way as to facilitate the changes in complementation listed in (a–d) above and exemplified in (47) and (48). Of course, this brief sketch has left many questions open, but it serves to illustrate in general terms how the principles-and-parameters approach to diachronic syntax can account for this kind of change.

One traditional and often repeated view is that clausal subordination, or hypotaxis, is a relatively recent reanalysis of parataxis, or clause-chaining (see for example Ernout and Thomas (1993: 291)). This idea has a long history, going back at least to Schlegel (1808) (see Harris and Campbell (1995: 25–7, 282ff.)). However, the claim that earlier stages of certain languages may have lacked subordination altogether violates the **uniformitarian hypothesis**, the idea that all languages at all times reflect the same basic UG, and so cannot be taken seriously in the approach adopted here. In fact, Harris and Campbell (1995: 282ff.) provide good arguments against this idea. Their most incisive criticism runs as follows: '[e]ven if parataxis does develop into hypotaxis, in and of itself this does not tell us how hypotaxis, true subordination, developed' (1995: 286). So I conclude that the traditional parataxis-to-hypotaxis idea should be abandoned, as it is conceptually problematic and in practice unrevealing.

On the other hand, it is quite plausible that a language may lack finite clausal subordination of the familiar type exemplified by English *that*-clauses and Romance *que/che*-clauses. In fact, Classical Latin was such a language, if what we have said about *quod*-clauses here is correct. Moreover, it is very likely that Classical Latin was in this respect typical of the older Indo-European languages; on this point, see in particular Kiparsky (1995). Turkish is an example of a language in which the familiar pattern of finite complementation plays a fairly marginal role: complementation is typically expressed by various kinds of nominalization (Kornfilt (1997: 45ff.)). In fact, the analysis of the development of Romance complementation sketched above implies that finite complementation is a parametric option, and the synchronic and diachronic evidence is that this is basically correct. Of course, it is entirely likely that a notion such as 'finite complement clause' is too coarse-grained, and would need to be replaced by something more abstract.

Related to the traditional parataxis-to-hypotaxis idea is the notion that adjunct clauses of various kinds may be reanalysed as complements. This idea is discussed in Harris and Campbell (1995: 287ff.), Kiparsky (1995); Roberts and Roussou (2003: 110ff.). Each of these approaches postulates a reanalysis roughly along the following lines:

- (64) I think [_{nominal} that] [_{clause} the world is round] >
 I think [_{clause} [C that] [the world is round]]

Reanalysis along these lines seems to have taken place in Germanic (Harris and Campbell (1995: 287–8); Kiparsky (1995); and Roberts and Roussou (2003: 116–20)) and Greek (Roberts and Roussou 2003: 120–1). Something like this may have happened at an early stage of the development of *quod*. The fact that finite complementizers very often develop from relative or demonstrative pronouns is clearly consistent with this. We can also note that the development of prepositional complementizers may involve a reanalysis partially similar to those seen in (61) and (64), whereby the sequence P+DP is reanalysed as C+TP (see Haspelmath (1989), Harris and Campbell (1995: 293)).

We have seen in this section that rather complex and pervasive changes in clausal complementation can be linked to two relatively simple but rather abstract parameter changes and their associated reanalyses. I have illustrated this with one case: the changes from Latin to Romance. However, there is nothing particularly unusual in these developments, and they are representative of the kinds of changes which can take place in complementation systems. If so, then what we have seen here is an illustration of how changes in the complementation system can be handled in terms of parametric change affecting the category C.

2.5. Word-order change: OV > VO in English

2.5.1. Introduction

In this section I will focus simply in the alternation between OV and VO orders, i.e. I will restrict attention to parameters F1 and F3 as defined in §1.6.1. Concerning parameter F1, we have already observed several times that Old English showed OV word order in embedded clauses. The following examples, which by now may be familiar, illustrate this:

- (65) a. ... þæt ic þas boc of Ledenum gereordre to Engliscre spræce **awende**.
 ... that I this book from Latin language to English tongue translate
 ‘... that I translate this book from the Latin language to the English tongue.’
 (AHTH, I, pref, 6; van Kemenade 1987: 16)

- b. ... þæt he **his stefne** up **ahof**.
... that he his voice up raised
'... that he raised up his voice.'
(*Bede* 154.28)
- c. ... forþon of Breotone **nædran** on scippe **lædde wæron**.
... because from Britain adders on ships brought were
'... because vipers were brought on ships from Britain.'
(*Bede* 30.1–2; Pintzuk 1991: 117)

We saw in §1.3.2 that OE had verb-second order in main clauses, and so we do not expect to find overt OV order in such clauses, unless of course the object is fronted to first position. The position of the auxiliary in (65c) indicates a further pattern which has changed since OE: the finite auxiliary in a subordinate clause followed the non-finite verb. This fact can be related to OV order in terms of parameter F3 of §1.6, which I repeat here:²⁷

(66) F3. Does the structural complement of V/T precede or follow V/T?

It appears, then, that F3 had the value PRECEDE in OE, while of course it has the value FOLLOW in NE. It has therefore changed in the course of the history of English. The goal of this section is to investigate in more detail what this assertion may mean, whether it is correct, and, if it is incorrect, how examples like those in (65) are to be interpreted.

2.5.2. *Early typological approaches to word-order change*

The earliest approaches to word-order change were directly inspired by Greenberg's (1963) observations of implicational relations among word-order types. W. Lehmann (1973) made two important proposals in this connection. First, he argued that subjects, since they may be dropped in many languages and can be pleonastic in any language (as far as is known), are not 'primary elements' of the clause. This reduces the word-order types to two: OV and VO. Second, Lehmann proposed that, in typologically consistent OV languages, verbal modifiers appear to the right of V and nominal modifiers to the left of O; in consistent VO languages we find the opposite pattern.

²⁷ F3 might be reformulated in the light of the postulation of vP between T and VP, but I leave this possible complication aside for the moment. I will return to v's possible role in word-order change in §2.5.4 below.

Now, many, or probably most, languages are inconsistent in relation to this typology (cf. NE, which is VO with preverbal modifiers, consistent with the typology, but it also has prenominal adjectives and possessors, inconsistent with the typology). To account for this, Lehmann proposed that ‘when languages show patterns other than those expected, we may assume they are undergoing change’ (55). Applied to the history of English, then, we could observe that, throughout its history, English has been drifting from OV to VO. Presumably, ME represents the period in which this very determinant of word order was in transition from one type to the other. This statement, in Lehmann’s system, would automatically predict a shift from VAux to AuxV order at the same period. Lehmann’s system also predicts a shift from RelN to NRel order in the history of English (see §1.6 for an illustration of this typological property); this is correct, although one could not claim that ME was the transition period for this change, as NRel already predominates in OE.²⁸ It seems unsatisfactory to consider the historical persistence of such ‘mixed’ systems as simply a feature of diachronic transition from one type to another. By this criterion, English has been in transition throughout its entire history, probably since Proto-Germanic. If F6 of §1.6.1 is a true parameter, then the same point could of course be made in relation to a putative change in its value from PRECEDE to FOLLOW.

Vennemann (1974) also advocated reducing Greenberg’s word-order types to OV and VO, leaving subjects out of the picture. He develops the Natural Serialisation Principle (or NSP, originally proposed by Bartsch and Vennemann (1972: 136)), which requires Operators and Operands to be serialized in a consistent order – either Operator Operand, or Operand Operator – in any language. Since objects, along with adjectives, relative clauses, etc. are Operators and verbs and nouns Operands, the NSP predicts the correlations with OV and VO orders which we observed in §1.6.1, and of course also predicts diachronic correlations. As in the case of Lehmann’s approach, the difficulty is that we are led to regard ‘mixed’ systems as persisting over very long periods, and, correspondingly,

²⁸ See J. Hawkins (1983: 222), who states that Late Common Germanic was already NRel. The claim that the oldest attested IE languages were already NRel is in fact important for Lehmann’s reconstruction of IE relatives (W. Lehmann 1974: 25). NE also retains Adjective-Noun order rather than the predicted Noun-Adjective order although ME arguably had a greater incidence of NA order than does NE (J. Hawkins 1983: 258).

individual changes in parameters like those in F3 as taking place over similar periods (cf. Vennemann's (1974: 353) remark that 'a language may become fairly consistent within a type in about 5000 years' (for example, English)). As various authors, for example, Comrie (1989) and Song (2001) have pointed out, this casts doubt on the strength of word-order conformity as a causal factor in change. This point is succinctly summarized in the following remark by Song (2001: 304):

[T]ypological consistency must at the same time be considered to be strong and weak ... It must be weak enough to permit incongruous word order properties to be incorporated into typologically consistent languages in the first place and it must also be strong enough to remedy the resulting situation by bringing all remaining word order properties into line with the new ones.

Regarding the synchronic predictions, J. Hawkins (1983: 41) points out that up to 77 per cent of the languages in Greenberg's original thirty-language sample do not conform to the NSP, and observes that 'the NSP's predictions are both too strong – there are too many exceptions – and too weak – there are distinctions between attested and non-attested language types that it is failing to capture' (42).

Again we see the essential empirical inadequacy of this kind of approach. And once again, we must bear this in mind in relation to the status of the generalized head-complement parameter F6 of §1.6.1.

Lightfoot (1999: 207ff.) makes a different kind of criticism, observing that long-term changes of the type envisaged by Lehmann and Vennemann are incompatible with a view of grammar as a cognitive module in the sense advocated by Chomsky. (This idea and the justification for it were presented in the Introduction to Chapter 1.) His point is that if grammars are psychological entities, then they are properties of individuals; they are reinvented anew with each generation of children. Long-term diachronic drift would then, all things being equal, entail a kind of 'racial memory' (209) on the part of the children acquiring language in order to cause the drift to continue in a consistent direction. Notions such as racial memory have no place in modern scientific theories, and so, to the extent that an approach like Lehmann's or Vennemann's, combined with a generally Chomskyan view of the nature of language, entails such a thing, we must reject either the Lehmann–Vennemann account or the Chomskyan view of language. Lightfoot strongly advocates rejection of the former. We will reconsider this argument when we come to consider Sapir's (1921) notion of drift in §4.3.

Despite these general difficulties with the NSP, Vennemann (1974: 359) proposes an interesting analysis of OV > VO change. The central idea relies on Greenberg's Universal 41:

- (67) If in a language the verb follows both the nominal subject and nominal object in dominant order, the language almost always has a case system (Greenberg 1963: 96).

Vennemann's central idea is that 'as reductive phonological change weakens the S-O morphology, and does not develop some substitute S-O [subject-object distinguishing – IGR] morphology, the language becomes a VX language.' This would naturally link the loss of case morphology, one type of 'S-O morphology', with the change from OV to VO. This seems attractive as an account of this change in both English and in the development from Latin to Romance, as this change was accompanied by the loss of morphological case marking (on non-pronouns) in both languages. However, it is clear that the loss of case is neither necessary nor sufficient for OV > VO change. Greek and Icelandic have both undergone this change while retaining their case systems; this point was also made by Kiparsky (1996: 142). (See A. Taylor (1994) on Greek, and Hróarsdóttir (1999) on Icelandic, which we also mentioned in §1.6.2.) According to Comrie (1989: 214–15), the Baltic and the Slavonic languages may be similar. Conversely, Dutch has largely lost its case system and yet has remained underlyingly OV, according to mainstream generative analyses, beginning with Koster (1975). Finally, Comrie (*ibid.*: 214) points out that Proto-Niger-Congo has been reconstructed as an SOV language without case, and many languages have changed to SVO, still with no case; and so this is an example of OV changing to VO quite independently of the loss of case.

J. Hawkins (1983: 134) proposed Cross-Categorial Harmony (CCH) as a generalization over many of Greenberg's implicational universals, as well as a number of exceptions to them. J. Hawkins states it as follows: 'there is a quantifiable preference for the ratio of preposed to postposed operators within one phrasal category ... to generalize to the others.'

The term 'operator' is taken from Vennemann's work, and so may be understood as described above. It is important to note that the principle is stated as a preference, rather than as an absolute requirement, and so grammars tend to correspond to it but do not have to. Furthermore, the principle makes reference to phrasal categories, explicitly acknowledging that phrase structure plays a role in accounting for these correlations.

In fact, J. Hawkins (179ff.) adopts X'-theory as the prime structural explanation for the CCH. X'-theory is the theory of phrase structure a variant of which was summarized in (55) of Chapter 1. The central idea of X'-theory is that all categories conform to the same structural template, which I repeat here as (68):

- (68) a. [_{XP} YP [_{X'} X ...]] (YP is a specifier of XP)
b. [_{X'} X YP] (YP is the complement of X)

As J. Hawkins points out (183), the basic advantage of X'-theory is that, since it offers a category-neutral template for phrase structure, it is well-suited to the expression of cross-categorial generalizations like the CCH. We observed in §1.6.1 that a general head-complement ordering parameter of the kind given there as Parameter F6 – repeated here as (69) – would predict spectacular cross-categorial harmony:

- (69) F6. For all heads H, does the structural complement of H precede or follow H in overt order?

However, it is clear that, as it stands, such a parameter is subject to criticisms of the kind summarized above in relation to the early proposals of Lehmann and Vennemann. J. Hawkins, however, interprets the CCH as a preference (and note that Dryer (1992) formulates his BDT as a preference too; see §1.6.1). In fact, J. Hawkins suggests that the CCH may derive from a preference for relatively simple grammars, since 'the more similar the ordering of common constituents across phrasal categories at the relevant bar levels, the simpler are the word order rules of the grammar.' If this is correct, then F6 cannot be a single parameter; cross-categorial harmony must derive from some higher-order factor determining interactions among formally independent parameter values, perhaps a simplicity metric of some kind. We will return to this idea in §3.5.

2.5.3. *Generative accounts and directionality parameters*

van Kemenade (1987), to some extent developing ideas in Canale (1978) and Hiltunen (1983), influentially applied X'-theory to accounting for the OV > VO change in the history of English. A general assumption in syntactic theory at the time was that there was a level of syntactic representation, known as the base, where the X' template held in a 'pure' form. (This assumption has been dropped in minimalist versions of syntactic

theory.) This template was subject to manipulation through movement operations at later stages of the derivation (cf. the discussion of Move in Chapter 1, Box 1.1). We can thus speak of ‘underlyingly’ VO and OV languages, i.e. we may take a parameter such as F3 or F6 to hold in the base but to be to some extent obscured by the action of subsequent movement operations. The word-order parameters assumed in this type of theory clearly have a more abstract status than the word-order variants assumed by Lehmann, Vennemann, and J. Hawkins. Moreover, this approach allows for the possibility that surface orders may diverge from the underlying order. As long as this divergence is somehow accessible to language acquirers, then the underlying order can be maintained, i.e. the parameter controlling the base order can be set. If the divergence becomes too great, then acquirers may reset the parameter (the Transparency Principle might again be relevant here; §2.1.2).

In these terms, a parameter like F6 derives from the option of head-initial or head-final order within the category formed by the head and its complement in the base, prior to the operation of any movement rules. Correspondingly, a parameter like F3 derives from a similar option where the head X is specified for some set of categorial features. This order may thus be H – XP (head-initial) or XP – H (head-final). van Kemenade assumed the OE order to be XP – H for H=V; this corresponds to our parameter F3, since van Kemenade assumes that auxiliaries, which we take to be T-elements, are verbs with sentential complements.²⁹ van Kemenade shows how certain movement operations disguised this underlying OV order in various ways. One such operation is known as extraposition, which moves a range of complements to the right of the verb, as shown by examples like the following:

- (70) a. ... þæt ænig mon atellan mæge [ealne þone demm]
 ... that any man relate can all the misery
 ‘... that any man can relate all misery’
 (*Orosius* 52.6–7; Pintzuk 1991: 36)

²⁹ In fact, van Kemenade took this order to be the consequence of the direction of assignment of thematic roles. If thematic roles are assigned to the right, the head assigning those roles precedes the complement being assigned them. If the roles are assigned to the left, the complement precedes the head. This, however, causes the parameter only to apply to cases where the head takes the complement as its semantic argument. It seems, though, that F6 has a wider purview than this, as was discussed in §1.6.1.

- b. *Æfter ðisum gelamp þæt micel manncwealm becom [ofer þære
after this happened that great pestilence came over the
Romaniscan leode].*
Roman people
'Then it happened that a great plague came over the Roman people.'
(*AHTh*, II, 122, 15; van Kemenade 1987: 40)

In (70a) an object DP, and (70b) a PP, appears to the right of the finite verb in a subordinate clause. We thus say that the DP and PP are extraposed.

Now, PP-extrapolation as in (70b) is also found in Modern Dutch and German, but not DP-extrapolation of the kind seen in (70a). Starting with Stockwell (1977), it has been proposed that OE, especially in later periods, extended the incidence of DP-extrapolation to a wider range of DPs than Dutch or German, in particular to 'light' DPs. Pintzuk and Kroch (1989) showed that in early OE (in the eighth-century epic poem *Beowulf*) only prosodically heavy DPs were postverbal in subordinate clauses and these were preceded by a metrical break. In later OE prose, on the other hand, this is clearly not the case; see Fischer *et al.* (2000: 148–9) for discussion. Examples like (71) illustrate this:

- (71) *Þu hafast gecoren [DP þone wer].*
thou hast chosen the man
(*ApT* 34.23; Fischer *et al.* 2000: 148)

Here, just the light DP *þone wer* is extraposed, and there is no evidence for a prosodic break after the participle *gecoren*. van Kemenade (1987: 41) concludes: 'It is quite possible then, that the phenomenon of extrapolation started off in early OE as the postposing of heavy constituents such as S [sentence/clause – IGR], PP and heavy NP's, and was extended later to include other constituents, light NP's, adverbials.'

A second factor was 'verb raising' and 'verb-projection raising'. These operations, the former found in Dutch and the latter in West Flemish and Swiss German dialects, derive orders in which the non-finite verb and possibly some of its complements appear to the right of the finite auxiliary. This is the opposite of the expected order in a language where V and T underlyingly follow their complements. For recent analyses of these phenomena in Modern West Germanic languages, see Hinterhölzl (1997) and Koopman and Szabolcsi (2000). OE had both operations, as the following examples illustrate:

- (72) a. *ðæt he Saul ne dorste ofslean* (verb raising)
that he Saul not dared murder

‘that he didn’t dare to murder Saul’

(*CP*, 199, 2; van Kemenade 1987: 59)

- b. þæt he **mehte his feorh generian** (verb-projection raising)

that he could his property save

‘so that he could save his property’

(*Oros*, 48, 18; van Kemenade 1987: 59)

van Kemenade (1987: 177) points out that the underlying order of OE was ‘not easily retrievable from surface patterns’ owing to the surface orders created by extraposition and verb(-projection) raising.³⁰ This was a major factor leading to the change in the parameter governing the underlying order. As a consequence ‘the underlying SOV order changed to SVO. This change was completed around 1200’ (van Kemenade 1987: 177). A similar idea is proposed by Stockwell (1977), while Stockwell and Minkova (1991) suggest that main-clause V2, which gave rise to many surface VO orders, may have influenced the acquisition of subordinate OV order, and once this became VO it in turn influenced main-clause order.

Lightfoot (1991) makes a different proposal regarding the relation between word-order change in main and embedded clauses. He proposes a broadly similar account to van Kemenade’s in that movement rules obscure the underlying order in such a way as to ultimately lead to a change in the value of the parameter determining the underlying order. He assumes a parameter determining underlying word order of the same general kind as that assumed by van Kemenade (his (6b, 42)). However, he assumes that language acquirers only have access to main clauses as trigger experience. They are ‘degree-0 learners’ in Lightfoot’s terminology, meaning that they can only access material which involves no clausal embedding. In OE main clauses, as we have seen (see §1.3.2), the finite verb appeared in second position. Hence the underlying OV order was obscured by the movement of the verb (to C, according to the standard analysis of V2 as summarized in

³⁰ van Kemenade also assumes that the parameter-settings responsible for OE word order were inherently marked. This is because she assumes that Nominative and Accusative Case are assigned from left to right while, as we mentioned in note 29, thematic roles are assigned from right to left. In current work, the earlier notion of Case-assignment is subsumed under the Agree relation introduced in §1.4.1, as we saw in §2.3. For this reason I leave this aspect of van Kemenade’s account aside here. The idea that extraposition and verb(-projection) raising would have obscured the underlying OV parameter value still holds, independently of these details of van Kemenade’s analysis.

§1.3). Of course, the same situation obtains in Modern Dutch and German. However, Lightfoot argues (52ff.) that '[v]erb-second languages typically have unembedded "signposts" indicating the movement site of the verb'. These include the final position of the particle where the verb is a particle verb (these are known as 'separable prefixes' in many pedagogical grammars of German), the position of the non-finite main verb where there is a finite auxiliary, or auxiliary-like, verb (i.e. where a single clause contains two verbal elements), and the position of 'verbal specifiers such as negatives and certain closed-class adverbials' which must, he assumes, be merged immediately to the left of VP. The linear separation of the finite verb from these elements in V2 clauses is illustrated by the following Dutch examples:

- (73) a. Jan **belt** de hoogleraar **op**.
 John calls the professor up.
 b. Jan **moet** de hoogleraar **opbellen**.
 John must the professor up-call
 'John must call the professor up.'
 c. Jan **belt** de hoogleraar **niet op**.
 John calls the professor not up
 'John doesn't call the professor up.'

This linear separation of V from its complement triggers V-movement, and allows the degree-0 learner to postulate the underlying OV word order. This is the situation in Modern Dutch and German.

Furthermore, Dutch and German both allow main-clause infinitives with a particular illocutionary force; (74) for example is a rhetorical question:

- (74) Ik de vuilnisbak buiten zetten? Nooit.
 I the garbage-can outside put? Never
 'Me put the garbage can outside? Never.'

In OE, however, at least two of the three 'signposts' showing linear separation of the finite verb and its complement are either absent or unclear, while the third has, according to Lightfoot, an uncertain status. The clearest observation is that OE negation involved the preverbal proclitic *ne*, rather than an adverbial element appearing to the left of VP as in Dutch and German. As a proclitic, *ne* is always directly adjacent and to the left of the finite verb:

- (75) Ne geseah ic næfre ða burh.
 not saw I never the city

'I never saw the city.'

(Ælfric, *Homilies* I.572.3; Lightfoot 1991: 62)

So the position of negation in OE does not function as a 'signpost' for the base position of the verb.

Second, particles were often fronted along with the verb in second position in OE, unlike their Dutch and German counterparts:

- (76) Stephanus **up-astah** þurh his blod gewuldorbeagod.
 S. up-rose through his blood glory-crowned
 (*Homilies of the Anglo-Saxon Church* I, 56; Lightfoot 1991: 61)

Regarding the position and status of non-finite verbs, Lightfoot (62) suggests that the relevant structures did not exist in OE. Thus, Lightfoot has grounds for asserting that the 'signposts' for underlying OV order in main clauses were less clear in OE than in Modern Dutch or German. Since, by the degree-0 hypothesis, language acquirers have no access to potential triggers in embedded clauses, the relatively systematic OV order of embedded clauses was not relevant to determining the value of the parameter.

On the other hand, OE did allow finite clauses with the verb in final position, in particular in the second conjunct of coordinate constructions, as in:

- (77) and his eagan **astungon**
 and his eyes (they) put out
 (Parker, *Anglo-Saxon Chronicle* 797; Lightfoot 1991: 58)

Such constructions gradually become less frequent during OE. Then 'as matrix instances of object-verb diminished to a certain point, underlying object-verb order became unlearnable and the verb-order parameter came to be set differently' (67). The indications are that this was an abrupt change in word order in embedded clauses in the twelfth century, as originally shown by Canale (1978). (We will see directly that this factual claim regarding both the date and suddenness of the change has been challenged.)

Whatever the merits of the idea that children are degree-0 learners, Lightfoot's account illustrates in a rather different way from van Kemenade's how a parameter determining underlying word order can change owing to that word order being distorted on the surface in a crucial way as the result of a movement operation.

Both Lightfoot and van Kemenade date the change in the relevant parameter to approximately the twelfth century. However, there are some difficulties with this. As van Kemenade (1987: 178) says, ‘the older word order did not, of course, become immediately ungrammatical . . . For a long time we continue to find OV structures, but . . . these were not firm enough in the language environment to trigger the older, marked situation.’ Similarly, Fischer *et al.* (2000: 162) point out that ‘[i]t is only after about 1300 that clauses with VO order begin to vastly outnumber those with OV order’.

They give, among others, the following late example of OV order, from Chaucer (late fourteenth century):

- (78) I may **my persone and myn hous so kepen and deffenden.**
 ‘I can keep and defend myself and my house in such a way.’
 (Chaucer *Melibee* 1334; Fischer *et al.* 2000: 163)

Foster and van der Wurff (1997) give the following ratios of VO to OV orders in poetry at fifty-year intervals in late ME: 2 (1350), 5 (1400), 13 (1450); in prose 4, 22 and 160. Clearly, some account must be given of the persistence of the OV orders into later ME.

Conversely, as already mentioned, OE shows a wider range of possible word orders, in both main and embedded clauses, than do Modern Dutch and German. Of particular importance in this connection are subordinate clauses with elements following a non-finite verb which are known not to appear to the right of such verbs in any Modern Germanic language: particles, pronouns, and light adverbs. These orders are illustrated in (79):

- (79) a. He wolde **adræfan ut** anne æpeling.
 he would drive out a prince
 ‘He would drive out a prince.’
 (*ChronB (T)* 82.18–19; Pintzuk 1991: 163)
- b. swa þæt hy **asettan him** upp on ænne sið.
 so that they transported themselves inland on one journey
 ‘so that they transported themselves inland in one journey.’
 (*ChronA* 132.19 (1001); Pintzuk 1993: 17)
- c. Þæt Martinus **come þa** into þære byrig.
 that Martin came then into the town
 ‘that Martin then came into the town.’
 (*ÆLS* 31.490–491; Pintzuk 1993: 17)

These orders look very similar to those of NE, and are usually interpreted as being an instantiation of the innovative order, in part because an extra-position analysis would either have to involve extraposition of particles,

pronouns or light adverbs, not usually thought to be possible, or of the implausible constituents consisting of these elements and the following material; *ut anne æbeling, him upp on ænne sið, þa into þære byrig* in the examples in (79), for example. If this is literally true, then OE must have already allowed head-initial orders. Pintzuk (1991; 2002) develops this idea by proposing that OE had a ‘double base’: both the OV and the VO value of the relevant parameter were allowed, giving rise to two distinct grammars through the OE period. This approach can account elegantly for some of the variation in word order (see Pintzuk (1991: 367ff.) for discussion), and we will return to the notion of grammars in competition in §4.1, §4.2, and §5.2.³¹ However, it is not clear what caused one of the two grammars (the OV one, in the case of the history of English) to fall out of use at the time it did. The notion of change in parameter values is not useful here. What appears to be clear, however, is that the transition from the old OV system to the new VO one was not as abrupt as van Kemenade and, in particular, Lightfoot, imply.

An important feature of the OV orders which appear in later ME is that they show a growing preponderance of quantified or negative objects. Kroch and A. Taylor (2000) show, by comparing a group of early thirteenth-century texts with a group of Late Middle English texts, that quantified objects appeared preverbally at both periods. In the later period, however, OV order was all but confined to quantified and negative objects. This conclusion is supported by data from fifteenth-century correspondence reported in Moerenhut and van der Wurff (2000) and Ingham (2001; 2002).³² Fischer *et al.* (2000: 163) state that fourteenth-century English continued to allow OV with non-quantified objects, but that fifteenth-century English

³¹ An interesting variant of this approach, where a major restriction is imposed on which categories can show parametric variation in head-complement order, and therefore coexisting word-order patterns, is developed by Fuß and Trips (2002). We will briefly consider Fuß and Trips’ proposal in §4.2.

³² Ingham (2001) relates the restriction to quantified/negative preverbal objects in fifteenth-century English to the constructions like that in (i), with the expletive *there* and a negative subject:

- (i) Ther shal no thing hurte hym.
(*PL* 209, 12; Ingham 2001: 23)

This is a further example of a transitive expletive construction, of the type discussed in §1.3.1.3.

only allowed OV where O was negative or quantified, or where there was an empty subject, as in a co-ordinate or relative clause. (80a,b) are examples from later ME of OV order with a quantified object, and (80c) illustrates OV order with an empty subject:

- (80) a. *Bei schuld no meyhir haue.* (negative object)
 ‘They were not allowed to have a mayor.’
 (Capgrave *Chronicles* 62.23; Fischer *et al.* 2000: 163)
- b. *He haþ on vs mercy, for he may al þynge do* (quantified object).
 ‘He has mercy on us, for he can do everything.’
 (*Barlam* 2740; van der Wurff 1999: 8)
- c. *alle þat þis writinge redder or heere*
 ‘all that will read or hear this writing’
 (*Sermon* 2250; Fischer *et al.* 2000: 163)

It seems, then, that OV order with quantified and negative objects was lost later than OV order with non-quantified and non-negative objects. In fact, Pintzuk (2002: 295–7) suggests that OV orders with quantified objects had a different status from those with non-quantified objects as early as OE. If so, then it is not surprising that the two types of OV order may have been lost at different times. In this connection, we note that the Greenbergian categories OV and VO are not sufficiently fine-grained to account for the observations that have been made. Finally, van der Wurff and Foster (1997) observe that surface OV disappears completely from prose writings during the sixteenth century.

Earlier, we criticized the Lehmann–Vennemann kind of approach to word-order change, in part because it leads to the conclusion that the change takes place over too long a period. What we have seen above suggests that an approach of the sort advocated by van Kemenade and Lightfoot has the change taking place too quickly; both before and after the alleged turning point, which they situate in the twelfth century, we find, respectively, the innovative order (cf. Pintzuk’s evidence from OE, of the kind in (79)) and the conservative order (it is clear in particular from Fischer *et al.* (2000: 163) that OV with a non-quantified object was found in the fourteenth century, as illustrated by (78)). Moreover, neither the gradualist Lehmann–Vennemann approach nor the catastrophist van Kemenade–Lightfoot approach can account satisfactorily for differences between early and late ME regarding negative and quantified objects in OV order. Where does this leave the idea that parameter F3 changed in ME?

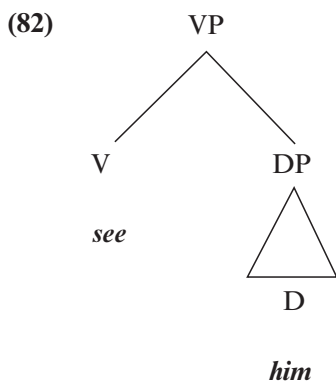
2.5.4. 'Antisymmetric' approaches to word-order change

We have several options in dealing with this situation. One option is to retract any attempt at cross-categorical generalization and revert to the position that the relative order of complement and head is to be restated for each category of head. As mentioned in §1.6.1, in connection with certain difficulties posed by German and Dutch for approaches to cross-categorical harmony, this would effectively make any implicational generalizations about word order of a synchronic or diachronic nature appear to be an accident.

Let us instead consider a different technical implementation of the parameters governing word-order variation. Kayne (1994) proposes the Linear Correspondence Axiom (LCA) as a principle of phrase structure. This can be stated as follows (this is an informal, simplified statement; for the original, see Kayne (1994: 5–6)):

- (81) A terminal node α precedes another terminal node β , if and only if α asymmetrically c-commands β .

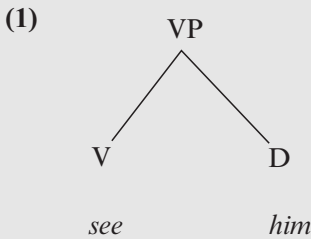
Terminal nodes are nodes which do not dominate anything, while non-terminals dominate something. We can see the implication of (81) for word order if we consider a simple verb-complement structure like (82), where for simplicity we assume that pronouns such as *him* are Ds:



Here V asymmetrically c-commands D and so, by the LCA, *see* must precede *him*. Given the LCA, there is no possibility of parametric variation in underlying head-complement order of the type assumed in particular by van Kemenade (1987) and Lightfoot (1991). Instead, the natural assumption (although not the only logically possible one) is that all languages are underlyingly VO, and OV orders are derived by leftward-movement of

BOX 2.1 Merge and the LCA

Strictly speaking, the presentation of the LCA here is not compatible with the structure-building operation Merge. In the Introduction, I presented Merge as the operation which ‘combines two syntactic elements (in the simplest case, two words) into a more complex entity which consists of those two elements and its label; the label being determined by one of the two elements’ (4). Merge is thus an intrinsically binary operation. However, in (82) *him*_D does not appear to have merged with anything. The discrepancy is in part due to the fact that Kayne (1994) does not assume Merge. If we ‘correct’ the structure in (82) to bring it into line with Merge we have (1):

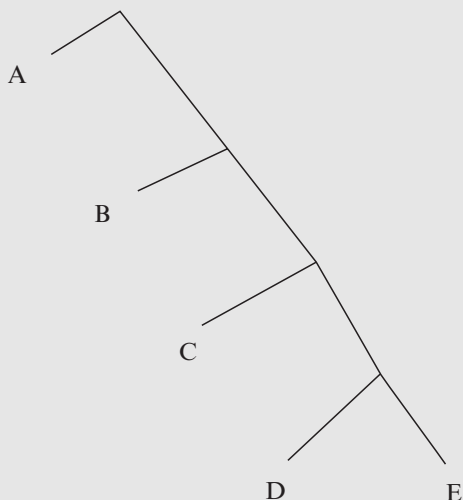


Here *him* is simultaneously maximal, in that it is immediately contained in a different category, and minimal, in that it does not itself contain anything. More generally, if we wish to make (82), and hence the LCA, compatible with the symmetrical nature of Merge as we have defined it we will always run into problems with the most deeply embedded category; the right branch of all higher categories is recursive and hence there is an asymmetric relation between the terminals which permits the LCA to determine linear order. This can be seen from the abstract phrase marker in (2).

The LCA defines the linear order $A > B > C > \{D, E\}$, but cannot order D and E. Since all trees must terminate with a non-recursive right branch, the problem of ordering the terminal on this branch with that on its sister will always arise.

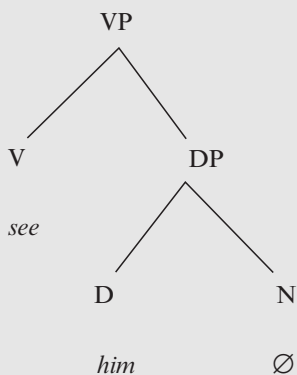
One way to solve this ‘rightmost-branch problem’, suggested by Chomsky (1995: 337), is to require that one of the elements merged at the most deeply embedded level is phonologically null. This is legitimate

(2)



since ‘there is no reason for the LCA to order an element that will disappear at PF’ (Chomsky 1995: 337). In this case, we can assume an empty N-element is merged with D, perhaps as a consequence of the inherent nature of D, so this would give (3) instead of (1):

(3)



Cardinaletti and Starke (1999) present a general analysis of the internal structure of pronominal DPs. It is not clear whether this proposal can provide a general solution to the rightmost-branch problem, however.

objects. This idea extends to all cases where a complement precedes its head in surface order; it must have been fronted to that position from an underlying post-head position determined by the LCA.

When I introduced the idea of movement in §1.3.1, I stated that movement was a matter of purely arbitrary variation among grammatical systems. In fact, Chomsky (2000; 2001) proposes that movement depends on Agree. A precondition for movement to relate two positions is that the two positions be in a Probe-Goal relation. Movement takes place where the Probe, in addition to having uninterpretable morphosyntactic features of some kind (see §1.4.1), also has an extra property which causes the Goal to undergo ‘second Merge’ (see Chapter 1, Box 1.1). Let us call this property ‘attraction’ of the Goal. Whether a given category is an attractor in this sense indeed appears to be a matter of arbitrary variation, although it is thought that only functional categories can be attractors. Following Chomsky (2000; 2001), I will indicate this attraction property with an Extended Projection Principle, or EPP, feature. (We mentioned the earlier conception of the EPP in note 17 of Chapter 1.) Where a Probe P is an attractor I will write it as P[+EPP]. Since many parameters involve the presence or absence of movement (for example, those connected to verb-movement discussed in §1.3), whether or not a Probe has an EPP feature is an important aspect of parametric variation; I will return to this point in §3.5.

Given this view of movement along with the interpretation of Kayne’s proposals just sketched, OV systems differ from VO systems in that some Probe P to the left of and structurally ‘higher’ than VP attracts the object, i.e. it enters the Agree relation with the object and triggers object-movement, so P would have an uninterpretable feature of some kind. The obvious candidate for P is v , since we saw in §2.3 that this element has uninterpretable φ -features which probe the direct object and allow the object’s ACC-feature to be deleted. OV order arises when v^* has an EPP feature. This gives the following derived structure:

(83) [_{v*P} DP-Obj v^* [u φ , EPP] [_{VP} V (DP-Obj)]]

The change from OV to VO must then be seen as the loss of the trigger for movement, i.e. the loss of v^* ’s EPP feature. Using different technical devices, an approach to word-order change in English of this kind was first proposed in Kiparsky (1996: 152) and developed in Roberts (1997). It has also been adopted by van der Wurff (1997; 1999); Fischer *et al.* (2000); and Ingham (2001; 2002), and by Hróarsdóttir (1999; 2000) for word-order change in Icelandic.

The approach, which we will refer to as the ‘antisymmetric approach’ since the LCA requires phrase structure to be antisymmetric, has a number of advantages. One conceptual advantage is that it eliminates the possibility of changes in the operation of Merge itself; only Move and Agree may vary and these are completely conditioned by the feature-content of functional heads. This is advantageous in that the range of formal options that the child must consider in setting parameters is limited. Indeed, we can continue to maintain that all parametric variation concerns Agree or Move; Merge is invariant. (We return to this point in §3.5.) Limiting the options of formal variation in this way is a good thing, in that it takes us a small step further towards reconciling poverty-of-the-stimulus considerations with the attested variation in grammatical systems (cf. the discussion in §1.1). This point has an important corollary when we consider change: it implies that changes must always concern Agree or Move, a point I will return to in §3.4 and §3.5. In this connection it is interesting to note Kiparsky’s (1996: 140) comment that ‘OV base order is commonly replaced by VO, whereas the reverse development is quite rare’. If OV order is derived by movement, as the antisymmetric approach implies, and if the loss of movement is a natural kind of syntactic change, then we understand why this is so.

Second, the antisymmetric approach to word-order change makes possible a more fine-grained empirical analysis, as movements can be selectively triggered. For example, we might claim that until around 1400, following Fischer *et al.*, v^* triggered movement of objects generally, much as schematized in (83), while in fifteenth-century English v^* only triggered movement of quantificational or negative objects, in virtue of having one of the two specifications $[u\varphi]$ or $[u\varphi, uOp, EPP]$. Since we have already seen that the Op feature may enter into Agree relations in our discussion of polarity items like NE *any* in §1.4.1, this suggestion has some independent plausibility. Of course, it is also possible that quantificational and negative objects were attracted by some other category than v , either from the fifteenth century (as suggested by van der Wurff (1997; 1999) and Ingham (2001; 2002)), or perhaps through ME and even in OE, as argued respectively by Kroch and Taylor (2000) and Pintzuk (2002). In that case v would have been $[u\varphi, EPP]$ up to c1400 and the other category would have been $[uOp, EPP]$ all along. Whichever of these analyses turns out to be correct, we can see that the antisymmetric approach, while conceptually more restrictive than the approach considered earlier, is also more flexible.

A further empirical advantage of the antisymmetric approach concerns the adjacency of the verb and direct object. To quote Kiparsky (1996: 173):

‘[r]igid VO languages ... require adjacency of verb and object ... whereas rigid OV languages ... allow adverbs to intervene freely’. We can observe this difference in the history of English: the examples of OE and ME OV order in (65) and (78) all have material intervening between the object and the verb; on the other hand, in NE the verb must be adjacent to the direct object (as mentioned in §1.3.1). If OV order is derived by leftward-movement of the object as in (83), then it is quite conceivable that adverbial and other material may intervene between the target of this movement and VP. VO order, on the other hand, does not involve object-movement, and so, to the extent that the verb does not move, the adjacency created by merging these two elements will be undisturbed. This difference between OV and VO systems cannot be so readily captured if we assume that the relevant parameter concerns the merged order of object and verb.

What I have said about the antisymmetric approach so far might lead one to consider that it is potentially so fine-grained that it has no hope of capturing larger-scale implicational relations such as those behind parameters like F3 or F6. But in fact this is not the case. If we ally the antisymmetric approach to the idea that potential movement triggers, i.e. various (sub)classes of functional heads, tend to trigger or fail to trigger movement harmonically, then we can in fact begin to understand the implicational relations that we have seen. Thus, parameter F3 would be the option of leftward-movement of the complements of V and T, and parameter F6 would be the option of leftward-movement of many different types of complements. The tendency for ‘head-initial’ and ‘head-final’ patterns to hold across categories, as revealed by the typological work of Greenberg, J. Hawkins, and Dryer, would result from a preference for potential movement triggers to act together. We could perhaps restate J. Hawkins’ (1983) generalization of cross-categorial harmony in the context of Kaynian antisymmetry as follows:

- (84) There is a preference for the EPP feature of a functional head F to generalize to other functional heads G, H ...

The reason behind (84) may reside in a kind of ‘meta-parameter’ governing the parametric options of individual functional heads, perhaps in a notion of **markedness** of parametric systems or perhaps in the nature of the parameter-setting process, i.e. language acquisition. I will return to this question, which concerns the nature of the theory of parameters, in §3.4 and §3.5, and very tentatively propose a more precise version of (84). For now it suffices to note that the antisymmetric theory is both fine-grained enough to allow an

analysis of the differences between fourteenth and fifteenth-century English, and at the same time at least in principle to capture large-scale implicational relations. To quote Kiparsky once more, the antisymmetric theory ‘would then predict ... that the mixed system of head-complement relations of Germanic would become uniform.’ But Kiparsky goes on to point out, as we have already mentioned, that ‘OV commonly changes to VO but the converse does not happen.’ As things stand, what has been suggested here does not predict this; in §3.5, I will return to this point.

A potential problem with the antisymmetric approach is that it runs the risk of entailing a complication of the structure of the clause, as triggers and landing sites for movement need to be postulated. To the extent that these are postulated purely to account for leftward-moved complements, the approach is no better than the one which postulates parameters determining the linear order of merged elements; it simply shifts the locus of variation from Merge to Move. If, on the other hand, the movements and positions needed to derive head-final orders are independently required, then this point is not problematic.

A second problem concerns the specific proposal that OV order is derived by leftward-movement of objects. In true OV systems, all complements must precede V. At first sight, this may seem to imply that a whole host of movements and positions must be postulated in order to account for preverbal PPs, particles, adverbs, etc. Again, to the extent that the movements are postulated purely to derive the head-final surface order, the antisymmetric approach loses its advantage over other approaches.

An interesting way of handling this last problem has emerged in recent years. Following an initial proposal by Hinterhölzl (1997), various authors (Haegeman 2000; Hróarsdóttir 1999; Koopman and Szabolcsi 2000; Koster 2000; Biberauer 2003) have proposed what one might call a ‘massive movement’ analysis for West Germanic languages. The basic idea is that the Goal for leftward-movement may be contained in a larger category which is moved along with the Goal (thereby undergoing pied-piping) when EPP-driven movement takes place. So, for example, instead of the object alone being attracted by v^* 's EPP feature, as in (83), the entire VP might move. This would give the structure in (83’):

(83’) [_vP [_{VP} V DP-obj v^* _[u_φ, EPP]] (VP)]]

In this structure, all VP-internal material in addition to the direct object is moved to the left of v^* . But, also, as it stands, V moves too, and so OV

order is not derived. But suppose V independently raises to v^* (Chomsky (2001: 35) assumes that this is a separate operation from verb-movement to T, object-movement, etc.), and then the ‘remnant’ VP moves to SpecvP. This will give the derived structure in (83''):

(83'') [v^* P [v_P (V) DP-obj] $V+v^*_{[u\varphi, EPP]}$ (VP)]

Example (83'') gives the surface order where all V's complements precede V. Suppose we now iterate the pied-piping operation at the TP-level, i.e. we allow T to attract the entire vP to its Specifier (pied-piping the subject, with which T's $[u\varphi]$ features Agree). This gives (85):

(85) [$_{TP}$ [v^*P DP-subj [v^*P [v_P (V) DP-obj] $V+v^*_{[u\varphi, EPP]}$ (VP)]] T (vP)]

If T is the category containing an auxiliary, then this sequence of operations will give us the surface word order *Obj-V-Aux*, which is the usual order found in subordinate clauses in West Germanic. Also, all other verbal complements will appear to the left of V and Aux.

Let us illustrate this kind of derivation with (65c), which we repeat here:

(65c) ... forþon of Breotone **nædran** on scippe **lædde wæron**
 ... because from Britain adders on ships brought were
 ‘... because vipers were brought on ships from Britain’
 (Bede 30.1–2; Pintzuk 1991: 117)

The order in which the VP-constituents are merged is as in (86):³³

(86) [v_P lædde of Breotone nædran on scippe]

When VP is merged with v^* , V raises to v^* and VP to the Spec v^*P , giving (87):

(87) [v^*P [v_P (lædde) of Breotone nædran on scippe] [v^* lædde $v^*_{[+EPP]}$] (VP)]

Next, the auxiliary *wæron* is merged in T and v^*P moves to SpecTP. This gives the structure in (88):

(88) [$_{TP}$ [v^*P [v_P (lædde) of Breotone nædran on scippe] [v^* lædde v^*] (VP)]
 [$_T$ wæron] (vP)]

Although this derivation appears rather complex, as long as the massive movements can be motivated, the approach has all the advantages of the antisymmetric approach to word-order variation and change which we

³³ If Merge is binary, as stated in the Introduction, there must be further structure inside the VP. I leave this aside here. The important point is that all these elements, except V itself, move as a unit.

enumerated above.³⁴ In particular the loss of these massive movement operations involving vP and VP pied-piping plays a central role in the OV > VO change we have discussed here.

2.5.5. Conclusion

In this section, we have discussed three different approaches to word-order change: the typological approach advocated by Lehmann and Vennemann, the approach postulating variation in underlying head-complement order of van Kemenade and Lightfoot, and the ‘antisymmetric’ approach based on Kayne (1994). We have also mentioned the ‘grammars in competition’ idea, influentially advocated by Pintzuk (1991; 2002) and Kroch and A. Taylor (2000). We have also seen that the data are more complex than the simple statement that English changed from OV to VO might seem to imply, while this statement nevertheless contains an important kernel of truth (and there is no reason to doubt that the same applies to other languages which have undergone this change, some of them mentioned in §1.6.2). In particular, OV order and its implicational correlate VAux appear to have been incrementally lost in English, beginning probably in late OE, with the final disappearance of OV only taking place at the end of the fifteenth century. If language change is driven by language acquisition, this cannot be a single parameter change; instead we must view OV order as arising from the interaction of several parameters, which tend to act harmonically. It seems that the antisymmetric approach lends itself particularly well to this view of things, although it is not without problems.

From the foregoing discussion, we see that word-order change in English is somewhat more complex than previously thought, in that it involves several separate but related parameter changes (see Fischer *et al.* (2000: 172–3) for a clear statement of what the various stages may have been). On this view, there

³⁴ We may further note that ‘verb-projection raising’ order as in (72b) and ‘verb-raising’ order as in (72a) can be straightforwardly derived by lack of vP-movement and object-movement instead of VP-movement respectively. In other words, these orders reflect movement of DP only to SpecTP and SpecvP, rather than pied-piping of vP and VP. We thus have a rather natural way of accounting for the synchronic variation in OE as a stable option of pied-piping vP or VP vs. ‘stranding’, i.e. movement of the DP Goal alone. This idea is developed in Biberauer and Roberts (2005a). We return to the question of optionality in §4.1.4.

is no simple OV/VO parameter, as different types of OV order are derived by different operations at different periods. This conclusion considerably refines our notion of ‘word-order type’ both synchronically and diachronically, and entails a J. Hawkins-esque notion of cross-categorial harmony, formulated in terms of the association of an EPP feature with different categories. An important conceptual advantage of word-order typology is thus retained, while the descriptive inadequacies of the ‘traditional’ OV/VO opposition are replaced by a more fine-grained analysis.

2.6. Conclusion to Chapter 2

The goal of this chapter was to illustrate the power and the utility of the notion of parametric change by showing how most of the principal kinds of syntactic change which have been discussed in the literature can be reduced to this mechanism. I have tried to show that reanalysis, grammaticalization, as well as changes in argument structure, complementation and word order, can all be understood in these terms. From here on, I will take it that this is the case; although there is much in the preceding two chapters that is open to debate, I maintain that they together constitute support for the thesis that the key notion for an understanding of diachronic syntax is that of parameter change. Furthermore, I follow Lightfoot (1979; 1991; 1998) in taking parameter change to be driven by language acquisition.

The notion of parameter itself remains unformulated, although in this chapter we have introduced one or two important notions (notably P-expression and P-ambiguity). One of the principal goals of the next chapter is to arrive at a proper characterization of a parameter as a formal aspect of the theory of grammar.

Further reading

Reanalysis, abduction, and learnability

Andersen (1973) is the classic exposition of the concept of abductive reanalysis. The empirical focus of the article is not syntax but phonology: sound changes in various Czech dialects. The conceptual importance of abductive reanalysis for our general understanding of change is however

made very clear. **Timberlake (1977)** is a classic study of syntactic reanalysis, in which it is proposed that the effects of reanalysis may not manifest themselves in surface changes immediately. **Longobardi (2001)** deals with the development of the French preposition *chez* from the Latin noun *casa*. The Inertia Principle plays a major role in the analysis, and its nature and implications are discussed in some detail. We will look at the Inertia Principle in §3.2. **Bertolo (2001)** is a collection of articles all dealing with aspects of learnability in relation to principles-and-parameters theory. **Kroch (2000)** is an excellent survey of the issues and results in generative diachronic syntax.

V-to-T movement and the development of English auxiliaries

Warner (1983) is a thorough and highly critical review of Lightfoot (1979), calling into question many of the empirical claims made there concerning the historical development of English modal auxiliaries. **Warner (1993)** is a monograph on the development of the auxiliary system in general, with the analysis stated in terms of Head-Driven Phrase Structure Grammar (HPSG). **Denison (1985)** defends the idea that auxiliary *do* developed from an earlier raising/control verb, which had a bare-infinitive complement, an idea developed a little further in Roberts (1993a). **Roberts (1985)** reconsiders the reanalysis of the English modals as auxiliaries, first dealt with in Lightfoot (1979). This is also the first paper to clearly recognize that English has lost V-to-T movement and to attempt to relate this to impoverishment of verbal agreement inflection. **Lightfoot (2006)** takes up these points, following on from the discussion in Lightfoot (1999); this book also presents a recent restatement of Lightfoot's views on a range of matters. **Vikner (1997)** gives a clear and systematic statement of the correlation between V-to-T movement and agreement inflection, which he restricts to VO languages. **Alexiadou and Fanselow (2002)** argue that this correlation is not an aspect of grammar, but rather a contingent fact about diachrony. **Anderson (2002)** is another critique of the proposals in Vikner (1997); again, the thrust of the argument is that conditions directly relating agreement inflection to movement are somewhat implausible. **Bobaljik (2002)** also criticizes Vikner's proposals, mainly on empirical grounds. **Thráinsson (2003)** looks at ongoing change in Faroese regarding V-to-T movement and agreement morphology, arguing that these developments pose problems

for the proposed correlation between V-to-T movement and rich agreement. **Bobaljik and Thráinsson (1998)** propose a parameterized version of Pollock's (1989) split-Infl idea: some languages combine T and Agr features on a single head, others split them into two projections. The cue for the difference is, again, the richness of verbal agreement morphology.

The effects of the loss of dative case and the development of psychological predicates in English

Allen (1995) is a thorough and interesting discussion of the development of recipient passives and constructions involving psych verbs. It is arguably the most in-depth study of the syntactic effects of the loss of the morphological marking of inherent Case in English to date. **Bejar (2002)** looks at Allen's analysis of the changes in psych verbs from a minimalist perspective. **Eythórsson and Barðdal (2005)** survey the behaviour of 'quirky subjects' in a range of Germanic languages, and come to conclusions only slightly different from Allen's. **Lightfoot (1981)** was one of the earliest analyses of the effects of the loss of morphological dative case on English syntax. **Fischer and van der Leek (1983)** is in part a response to this, going into much greater empirical detail regarding the development of ME psych verbs. **Van der Gaaf (1904)** is the principal traditional study of the development of psych verbs in the history of English. **Belletti and Rizzi (1988)** is an influential analysis of psych verbs, mostly in Italian, in terms of government-binding theory. **Pesetsky (1994)** includes a very detailed study of NE psych verbs, breaking them up into a number of thematically-defined subclasses. **Baker, Johnson, and Roberts (1989)** is an influential analysis of passives using government-binding theory. **Collins (2005)** is the most thorough and interesting alternative account of passives, which develops an important idea concerning restrictions on the locality of movement in minimalism.

Grammaticalization

Bybee, Perkins, and Pagliuca (1994) is a major typologically-based survey of grammaticalization phenomena. **Heine et al. (1993)**; **Heine, Claudi, and Hünemeyer (1991)**; **Heine and Kuteva (2002)**; **Traugott and Heine (1991)**; and **Heine and Reh (1984)** are all important collections of materials on

grammaticalization, again looked at from a functional-typological perspective. **C. Lehmann (1986; 1995)** provides useful overviews of the phenomena. **Haspelmath (1989)** presents an account of the development of infinitival markers from purposive conjunctions, relying on the functional-typological notion of grammaticalization. **Batllori et al. (2005)** is a recent collection of articles adopting a formal, mostly minimalist, approach to different kinds of grammaticalization phenomena. **Meillet (1912)** is a general paper on grammatical change. It is noteworthy for the first recorded use of the term 'grammaticalization' and for the claim that this, along with analogy, are the only mechanisms of grammatical change. **Bopp (1816)** was one of the first major treatises on comparative Indo-European grammar. Some of the ideas put forward prefigure more recent ideas about grammaticalization. **Hopper and Traugott (2003)** is the main textbook on grammaticalization. Again the focus is largely functional-typological. **Simpson and Wu (2001)** is an interesting formal account of grammaticalization in various East Asian languages. **Wu (2000)** is a formal treatment of a number of cases of grammaticalization in the history of Chinese. **Roberts and Roussou (1999)** is an early version of the later monograph on grammaticalization, in which the formal, minimalist-based approach is proposed (Roberts and Roussou 2003; see Further reading to Chapter 1).

Word-order change in English

Foster and van der Wurff (1997) study OV vs. VO orders in Late ME, with some very revealing quantitative results. **Moerenhout and van der Wurff (2000)** is another partly quantitative analysis of the incidence of OV orders of various kinds in ME. **Van der Wurff (1997; 1999)** and **van der Wurff and Foster (1997)** further investigate details of ME word order, supporting an antisymmetric approach and arguing that object-movement in Late ME was restricted to certain types of object. **Ingham (2001; 2002)** looks at the nature of the preverbal object in Late ME OV orders, showing clearly the preference for negative or quantified objects in this order. **Kroch and Taylor (2000)** argue that quantified objects move to a designated position throughout the ME period. **Pintzuk and Kroch (1989)** is an important early study of OE word order, in which the authors demonstrate that postverbal objects in subordinate clauses in *Beowulf* are always preceded by a metrical pause. They argue that this is consistent with the idea that such objects are

extraposed. **Roberts (1997)** proposed an ‘antisymmetric’ VO analysis of OE word order, suggesting that this has the advantage of allowing us to see the change from surface OV to VO as the loss of leftward object-movement rather than as a reanalysis of the base. **Biberauer (2003)** proposes an account of synchronic variation in Modern Spoken Afrikaans which makes use of massive movement and pied-piping options as sketched in §2.5.4. **Biberauer and Roberts (2005a)** applies these ideas to word-order variation and change in the history of English. **Stockwell (1977)** and **Stockwell and Minkova (1991)** are further studies of OE word order and ME word-order change, the former being one of the earliest generative analyses of OE. **Canale (1978)** is an early study of word-order change in ME. **Hiltunen (1983)** is another important early study.

Other work on the history and varieties of English

Denison (1993) is a comprehensive review of nearly all the major work on the historical syntax of English available at the time, very coherently organized into topical sections and with extremely useful commentary and bibliography. **Henry (1995)** is a detailed study of a number of syntactic peculiarities of the English of Belfast. Some of them, including the Northern Pronoun Rule, are shared by other regional varieties of English deriving from Northumbrian OE. **C. Jones (1997)** is a comprehensive survey of Scots English, with much useful historical material. **Los (1998)** looks at the rise of the *to*-infinitive in ME, arguing that it did not derive directly from an OE purposive, but had an earlier origin, with its distribution being enlarged during ME as it replaced subjunctive *that*-clauses in a number of contexts. **Jespersen (1909–49)** is a classic survey of the historical grammar of English. It remains the most comprehensive work of its kind. **Visser (1963–73)** is a very large compendium of syntactic constructions from all periods of English. Before the advent of electronic corpora, this was an invaluable tool, and it remains useful today.

Germanic syntax

Hinterhölzl (1997) deals with verb-raising and verb-projection raising in dialects of German. This was one of the first studies in which a ‘massive

movement' analysis was proposed in order to account for this kind of phenomenon. **Koopman and Szabolcsi (2000)** develop and extend the massive-movement idea to a range of constructions including those involving preverbs in Hungarian. **Koster (2000)** invokes massive movement in the analysis of verb-raising in Dutch. **Wurmbrand (2001)** is an analysis of verb-raising and related restructuring phenomena in German, in which it is argued that these 'clause-union' phenomena involve reduced complements, probably vPs. **Kiparsky (1995)** presents an intriguing reconstruction of clausal subordination in Indo-European, and a proposal for the development of finite complementizers in Germanic which relates this to the rise of V2. **Koster (1975)** is a classic article in which it is shown that the most economical analysis of Dutch V2 clauses involves generating the verb in final position and raising it to C. **Sigurðsson (1989)** is an important study of case marking and non-finite clauses in Icelandic. The major result is the evidence that PRO can bear dative case, contrary to a central tenet of government-binding theory. **Jonas (2002)** looks at various aspects of the syntax of Norn, the North Germanic language spoken on the Shetland Islands until the eighteenth century.

Latin and Romance syntax

Bolkestein (1979) provides detailed arguments that the Latin accusative + infinitive construction is not the same as English-style Exceptional Case-Marking. **Cecchetto and Oniga (2001)** develop Bolkestein's analysis further, adding more data and greater sophistication. **Sihler (1995)** is an important and very thorough historical grammar of Latin and Greek, clearly demonstrating the relations between these languages and Indo-European. **Vincent (1988)** is a very useful survey of the structure of Latin, with a thorough discussion of how Latin clausal complementation differs from that of the Modern Romance languages. **Woodcock (1959)** is a traditional grammar of Latin. **Perlmutter (1978)** was the first to observe the systematic differences in behaviour between unaccusative and unergative intransitives in Italian. He provided a detailed analysis of these and related constructions in terms of relational grammar. **Burzio (1986)** was the first in-depth study of unaccusatives and related constructions in government-binding theory. It also deals with a very wide range of constructions from Italian and its dialects. **Levin and Rappaport-Hovav (1995)** is the most thorough study of

unaccusativity in English to date, featuring very insightful analyses of a number of subsystems of the English lexicon. **Guasti (1991)** presents an analysis of the complements of causative and perception verbs in French and Italian in terms of a late version of government-binding theory. **Calabrese (1993)** develops an analysis of control and raising in Salentino, a Southern Italian dialect almost entirely lacking in infinitives. **Ledgeway (1998)** similarly studies these and related phenomena in the dialects of Southern Calabria and North-East Sicily, which appear to entirely lack infinitives. **Ledgeway (2000)** is the most detailed study of the syntax of Southern Italian dialects to date. **Rohlf's (1969)** is the classic traditional description of Italian dialects. **M. Jones (1996)** is probably the most comprehensive survey of Sardinian syntax in terms of generative grammar to date. **Kayne (1983)** is a classic treatment of 'complex inversion' in French. **Rizzi and Roberts (1989)** reanalyse French complex inversion, exploiting the VP-internal subject hypothesis in order to account for the presence of two realizations of the subject in this construction. **Roberts (1993b)** is an analysis of clitics and inversion in Franco-Provençal Valdôtain, a variety of Franco-Provençal spoken in the Val d'Aoste in North-Western Italy. **Kayne (2000)** is a collection of articles on universals, Romance syntax and English syntax, with an introduction in which, among other things, Kayne argues that the number of distinct grammatical systems currently extant is probably greater than the human population. **Wheeler (1988)** is a thorough survey of the history and structure of Occitan.

The structure of DPs

Bernstein (1991; 2001) present a general analysis of the structure of nominals across languages using the DP hypothesis and N-to-D movement. **Longobardi (1994)** is another major study of nominal syntax in terms of the DP hypothesis, in which N-to-D movement plays a central role. **Ritter (1991)** also looks at nominals in terms of the DP hypothesis and N-to-D movement. **Zamparelli (1995)** is another treatment of the internal structure of nominals, arguing in particular, and partly on semantic grounds, for a distinct functional projection from D to house certain types of quantifier.

French phonology

Dell (1985) remains the main study of French phonology, in terms of ‘classic’ generative phonology. **Pagliano (2003)** is a recent treatment of liaison in French, showing that the /t/ which appears in *a-t-il* and similar contexts is epenthetic, while the /t/ that appears with inverted verbs in the 3pl, as in *ont-ils*, is an instance of liaison and such may involve an underlying /t/. **Tranel (1981)** is another major study of French phonology.

This page intentionally left blank

3

Acquisition, learnability, and syntactic change

Introduction	207	3.4. Markedness and complexity	251
3.1. First-language acquisition from a principles-and- parameters perspective	209	3.5. Parameter setting and change	266
3.2. The logical problem of language change	226	3.6. Conclusion to Chapter 3 Further reading	282 284
3.3. The changing trigger	236		

Introduction

Having by now established that parameter change can describe many instances and types of syntactic change, here we look at the deeper questions this conclusion raises. The purpose of this chapter is to explore the idea, introduced in §2.1, that parameter change is driven by the first-language acquisition process (this idea is pursued in Lightfoot (1979; 1991; 1998); see Croft (2000: 47–9, 119) for critical discussion), and thereby to illustrate how the study of syntactic change may be relevant for our understanding of the processes involved in first-language acquisition. One way of construing the idea that parameter change is driven by the first-language acquisition process is to think that a parameter

value changes because an innovative alternative is more ‘accessible’ to acquirers, thereby rendering the conservative value in effect ‘inaccessible’, or unlearnable. This view has two important consequences. First, to the extent that abductive reanalysis of the sort discussed in §2.1 is a symptom of an underlying parameter change, it can explain the pervasiveness of this kind of reanalysis in syntactic change. Second, it entails that language **learnability** is intimately connected to change – in fact learnability becomes the key to understanding syntactic change. This further tightens the connection between L1 acquisition of syntax and syntactic change. In fact, Niyogi (2004: 462) points out that ‘every theory of language acquisition also makes predictions about the nature of language change.’

We begin by looking in §3.1 at the current state of knowledge regarding first-language acquisition of syntax, basing our presentation fairly closely on the discussion in Guasti (2002). In §3.2 we consider what we call, following Clark and Roberts (1993), ‘the **logical problem of language change**’; we will see that this is closely related to, and maybe subsumes, the Regress Problem discussed in §2.1. In this context, we introduce a central idea for much of the later discussion: the idea that the language learner (or, in more technical terms, the parameter-setting device) is computationally conservative, obeying a kind of ‘least-effort’ constraint, i.e. a general preference for simplicity of representations, which we formulate along the lines of Roberts and Roussou (2003: 201). In §3.3 we try to get a picture of what kinds of external circumstances could cause a parameter change; this relates closely to the discussion of the logical problem of language change. Under this heading, we discuss in a preliminary way the possible role of language contact (although this will be the focus of Chapter 5), as well as the notion of cue introduced by Dresher (1999) and discussed at length in relation to syntactic change in Lightfoot (1999). We also discuss the role of morphological change in triggering syntactic change. §3.4 introduces the notion of markedness and relates it to the characterization of complexity/simplicity given in §3.2. Finally, in §3.5, we try to bring the strands of the discussion together in a general proposal for the form of parameters, how they are set and how they may change. This concludes the general discussion of parametric change as the mechanism of syntactic change, the remaining chapters being concerned with the wider implications of this view. But let us now begin at the beginning, i.e. with first-language acquisition.

3.1. First-language acquisition from a principles-and-parameters perspective

In this section I will try to rather sketchily summarize some aspects of the burgeoning recent literature on first-language (L1) acquisition of syntax. My focus will be on the major empirical observations and their implications for the thesis that parametric change is driven by the acquisition process. These are, first, that many important parameters appear to be set rather early in the acquisition process (see (2) below), and, second, that there are two phenomena of interest in children's early production: the so-called **root** or **optional infinitives** (see Radford (1990; 1996); Platzack (1992); Pierce (1992); Wexler (1992; 1994; 1999); Poeppel and Wexler (1993); Rizzi (1994); Haegeman (1995b); Hoekstra and Hyams (1998); Hyams (1996); Schütze (1997); Hamann and Plunkett (1998); see also Guasti (2002: 128ff.) and the references given there) and 'early null subjects' (see *inter alia* Hyams (1986; 1992); Bloom (1990); Valian (1990); Gerken (1991); Weissenborn (1992); Hyams and Wexler (1993); Rizzi (1994; 2000); Clahsen, Kursawe, and Penke (1995); Haegeman (1995a); Guasti (1996; 2000); and the papers in Friedemann and Rizzi (2000), as well as the references given in Guasti (2002, Chapter 5)). The purpose of our discussion is to indicate to what extent our understanding of the parameter-setting process has been furthered by this work, and to see if in principle any connection with a parameter-changing approach to syntactic change can be discerned.

Before looking at the phenomena which have been observed, however, we need to be clear about our general conception of first-language acquisition. In the introduction to Chapter 1, I presented and tried to justify Chomsky's claim that the human language faculty is a facet of human cognition, physically instantiated in the brain and, most importantly, genetically inherited as an aspect of the human genome. Under this conception of the language faculty, first-language acquisition can be characterized in the following terms:¹

¹ As we mentioned in note 1 of Chapter 1, modularity – the idea that the language faculty is a distinct system of the mind/brain – may not play a role in the minimalist conception of the language faculty. But the crucial point for the purpose of the discussion of first-language acquisition here is that even if the language faculty (in either the broad or the narrow sense as defined by Hauser,

The language faculty is a distinct system of the mind/brain, with an initial state S_0 common to the species ... and apparently unique to it in certain respects [footnote omitted]. Given appropriate experience, this faculty passes from the state S_0 to some relatively stable state S_S , which then undergoes only peripheral modification.

(Chomsky 1986: 25)

The initial state of language acquisition is nothing other than Universal Grammar, while the stable state is adult competence in a given language, which remains unaltered in essential respects from childhood on. The process of first-language acquisition is the process by which the language faculty ‘passes from the state S_0 to some relatively stable state S_S ’ in Chomsky’s formulation. We thus have the following schematic notions:

- (1) a. $UG = S_0$ (initial state);
- b. Adult competence = S_S (stable state);
- c. Stages of acquisition = $\langle S_0, \dots, S_{i>0}, \dots, S_{n>i}, S_{n+1<S}, \dots, S_S \rangle$.

Here, (1c) indicates the stages of first-language acquisition. These can be thought of as an ordered n -tuple of states of indeterminate, but certainly finite, number, occurring later than S_0 and earlier than S_S . They correspond neither to UG nor to the adult competence, but rather to what we can think of as an immature competence. (I will say more about this notion of immaturity below.)

How are the various states defined in relation to one another? To put it crudely, what does S_{n+1} have that S_n lacks in (1c)? As we have seen, the ‘innateness hypothesis’ claims that S_0 is determined wholly by the genome, independently of any experience. On the other hand, S_S differs from S_0 in that it is at least partly determined by experience of the linguistic environment (the primary linguistic data, or PLD): the PLD, among other things such as providing the vocabulary of the first language, in some way causes parameters to be set to determinate values. We can therefore assume that the various intermediate states are distinguished by having differing values of various parameters. Each S_n differs from S_{n+1} in one of two ways: either insufficient experience has been accumulated at S_n for setting certain

Chomsky, and Fitch (2002)) is in some sense ‘emergent’ and may certainly lack a single neurophysiological locus in brain architecture and a single phylogenetic source, we can nevertheless meaningfully distinguish the initial state of the system (or of the relevant subparts) in the newborn child from the modified state which is the stable, adult state. This process appears to be subject to a critical period; see §5.4 for some recent evidence for this.

parameters, or the overall system has not matured sufficiently at S_n to permit certain parameters or parameter values to be attained. This in turn could be due to different parameters coming ‘on line’ for acquisition at different times, to the system gradually maturing with respect to the kinds of PLD it is able to accommodate, or to interactions among already-set parameters, perhaps in accordance with a general notion of markedness (see §3.4 and §3.5). In any case, we can consider that the various intermediate states differ from one another in representing successively closer approximations to the adult system (S_S) in terms of the values of the parameters. To put it another way: if m parameters are set to the adult value at stage S_n then at least $m + 1$ parameters are set to the adult value at stage S_{n+1} .

The nature of the intermediate grammars has been studied fairly intensively in the past twenty years or so, beginning with the pioneering study in Hyams (1986). Guasti (2002) points out that between the ages of two and three years old, i.e. some time before linguistic maturity if this is characterized as the stable state, children know at least the following about the parameter values of the language they are in the process of acquiring (the references are given in Guasti (2002: 148, 185, 242)):

- (2) a. the value of the head direction parameter in their native language;
- b. the value of the V-to-T parameter in their native language;
- c. the value of the topic-drop and null-subject parameters;
- d. the value of the parameters governing question formation, the one governing overt movement or in-situ placement of the wh-element and the one regulating T-to-C movement (inversion).

To this we can add that Hamann, Rizzi, and Frauenfelder (1996) show that as soon as a French-acquiring child produces clitics, they are placed in the correct clitic position for French, even though there is much parametric variation in clitic-placement across languages, and, following Wexler (1998), verb second. Wexler (25) describes this general phenomenon as follows: ‘[b]asic parameters are set correctly at the earliest observable states, that is, from the time that the child enters the two-word stage around 18 months of age.’ He continues: ‘[q]uite possibly . . . children have set basic parameters . . . *before* the entry into the two-word stage’. This observation has become known as Very Early Parameter Setting, or VEPS.

Guasti also provides evidence that, also between the ages of two and three, children know the properties of unaccusative verbs (see §2.3.1).

Furthermore, by the age of four, children comprehend and produce passives based on actional verbs, although they have difficulty with passives of non-actional verbs (2002: 269); they have acquired most, but not all, of the principles governing the distribution and interpretation of anaphoric and other pronouns (the **binding theory**: see the textbooks mentioned in the Introduction for discussion of this aspect of syntactic theory) (2002: 310); they have also acquired the principles concerning the distinction between referential and **quantified expressions** (for example, *John* vs. *every boy*) and many aspects of the interpretation of quantified expressions (2002: 344); and, finally, that they have acquired many but not all aspects of the nature of the ‘control’ relation between a DP in a superordinate clause and the understood subject of a non-finite subordinate clause briefly alluded to in §2.4 (2002: 372).

Most of the parameters listed in (2) are familiar from Chapter 1. (2a) refers to word-order parameters, of which we identified several subtypes in the discussion in §1.6.1 (parameters F1–F6). The L1-acquisition literature has shown that children are sensitive to these parameters and that ‘from the onset of multiword utterances (or even earlier)’ (Guasti 2002: 103) they have correctly identified the relevant values for the ambient language. (2b) clearly refers to parameter B of §1.3.1. (2c) partly refers to parameter A of §1.2.1, although the notion of ‘topic-drop’ was not discussed there; I will return to this in the discussion of ‘early null subjects’ in L1 acquisition below. The second part of (2d) concerns parameter C, T-to-C movement, while the first part refers to parameter E of §1.5.

All of the parameters listed in (2) are important and salient for synchronic description, and Guasti’s summary of the L1-acquisition evidence shows that they are salient for language acquirers; these parameters are acquired early and correctly, or so it appears. As it stands this observation supports the poverty-of-the-stimulus argument, as Guasti implies (2002: 147); children are able to glean the values of these parameters from the PLD almost before they are able to produce multiword utterances. This strongly suggests a predisposition to the task, given the very young age of the children (multiword utterances normally begin between twenty and twenty-four months old; see Guasti (2002: 98)), the complexity of PLD, and the rather abstract nature of these parameters. However, we saw in Chapter 1, several of these parameters can be shown to have changed their values in the recorded history of various languages. This brings us face to face with what Clark and Roberts (1993: 299–300) termed the logical problem of

language change, which I will discuss in more detail in the next section. For the moment, it suffices to note both the tension between Guasti's conclusions and what we saw in Chapter 1, and that the obvious resolution of this tension must involve some intergenerational change in the nature of the PLD; some of the issues surrounding this conclusion were discussed in §2.1, and will be discussed further later in this chapter.

Returning to the discussion of the intermediate stages of acquisition as defined in (1c), it may seem that we have little evidence regarding the setting of the parameters in (2) beyond the fact that they are typically set early and accurately. Far from shedding light on a parameter-resetting approach to syntactic change, this appears to pose a problem, as just noted. Of course, it is entirely possible that parameter-resetting can take place between twelve and twenty-four months, but this will be very difficult to document on the basis of children's utterances prior to the two-word stage (which normally starts at around eighteen months); it is possible that some experimental methodology could be developed in order to ascertain this, but I am aware of none that has been developed at present. (Wexler (1998: 25, note 1) makes the same point.) But the grammar of the earliest stages of multiword production is still an immature one in the sense that it is subject to modification through further stages of acquisition. Guasti's conclusions as listed in and immediately below (2) appear to show that there is an intermediate, fairly early, state of the language faculty $S_{i>0}$ at which a number of important parameters have been set, but that there are nevertheless further stages of acquisition, and presumably therefore of parameter-setting, $S_{n>i}$, $S_{n+1<S}$, etc., remaining. (See note 5 below for a conjecture as to the difference between the two sets of parameters.) Since these are stages during which children produce multiword utterances, we have in principle better access to the parameter-setting process here than in the case of the parameters whose values are set earlier. Can anything relevant for our conception of syntactic change as parameter-resetting be gleaned from these later stages?

What would be the 'ideal scenario' for relating parameter setting in first-language acquisition to parameter change? The kind of case which could link the evidence of production based on immature grammars of the type discussed in the L1-acquisition literature to the questions relevant for syntactic change would have to have four properties. First, we would want to compare the acquisition of two closely related languages L and L' where it is known that L' is syntactically innovative in relation to L

(in that it is known that the common parent language of L and L' set some parameter P to value v_i , and that L has value v_i for P while L' has value $v_{j \neq i}$ for P). Second, we would observe that early production in L showed a tendency for strings which appear to express value v_j for P, in the sense of parameter expression introduced in §2.1, (8). Third, we would observe that this apparently aberrant production in L ceases when some further feature F of L is acquired, i.e. manifested in production, and, fourth, we would observe that F was lost from L' when P changed value. So we link the change in P, *both in acquisition and in diachrony*, to the presence/absence of F.

The two main phenomena in early production which have been looked at, root/optional infinitives and early null subjects, come close to instantiating the 'ideal scenario' as just described in relation to some of the parameters listed in (2). It is probable, however, that neither of them truly instantiates this scenario. Nevertheless, it is worth looking at them.

The essence of the root (or optional) infinitive phenomenon is that it 'consists of producing main clauses containing an infinitive verb, rather than a finite one' (Guasti 2002: 128). It is 'peculiar to the earliest multiword productions and lasts until about 3 years' (*ibid*). Some examples from various languages are given in (3) (taken from Guasti (2002: 128–9); sources for the examples are given there):

- (3) a. hun sove (Swedish)
she sleep-infin
- b. earst kleine boekje leze. (Dutch)
first little book read-infin
- c. pas manger la poupée (French)
not eat-infin the doll
- d. s[ch]okolade holen (German)
chocolate fetch-infin
- e. Papa have it. (English)

In each example, the verb has the form of an infinitive, despite the fact that these are all main clauses. For English, this implies that the verb has the bare-stem form: cf. *have* rather than *has* in (3e). In the other languages, the ending is recognizably that of the infinitive (for example, *-en* in Dutch and German, *-er* in French). Moreover, in some of the examples, the verb has the syntax of an infinitive: it follows the negative element *pas* in French in (3c), while finite verbs must precede *pas* (see §1.3.1), and it follows the

direct object in Dutch – (3b) – and German, (3d). Here the Dutch example is more telling, since this could not be a verb-second clause, given the presence of the adverb *earst* (*erst?*) ‘first’ in addition to the direct object.

The following constraints on root infinitives have been observed in the L1-acquisition literature:

- (4)
- a. Root infinitives do not occur in null-subject languages.
 - b. Root infinitives are not introduced by nonsubject XPs in V2 languages.
 - c. Root infinitives are incompatible with clitic and weak-pronoun subjects.
 - d. Root infinitives occur in declaratives, but not in *wh*-questions.
 - e. Root infinitives are incompatible with auxiliaries.

In connection with (4a), note that all the languages exemplified in (3) are standardly analysed as non-null-subject languages (see §1.2.1). (4b) strongly suggests that clauses containing root infinitives are not verb-second clauses, even in verb-second languages, despite being main clauses – see below. (4c) is fairly self-explanatory (recall the notions of clitic and weak pronoun from the discussion of the history of French in §1.2.2), and (4d, e) are straightforward.

Two principal types of analysis have been proposed for this phenomenon: the Tense-omission approach of Wexler (1994; 1999), and the clausal-truncation approach of Rizzi (1994). (Avrutin (1998) pursues a third option, observing that root infinitives occur in adult Russian under certain discourse conditions; it is therefore possible that the cases illustrated in (3) are straightforward instances of initial parameter-missetting; it is not clear, however, to what extent Russian and some related language might instantiate the ‘ideal scenario’ as described above). Both approaches rely, in different ways, on the idea that the T-position in main clauses optionally lacks some crucial property in the child grammars which produce root infinitives: for Wexler, the tense feature is not specified, while for Rizzi, all projections above VP may be ‘truncated’, i.e. simply not present, at this stage of grammar development; Rizzi (2005: 94–5) extends this possibility to the categories making up the ‘split CP’ he assumes. Both analyses can account for the absence of root infinitives in null-subject languages: for Wexler, this depends on specific assumptions about the relationship between the T-position and the agreement features characteristic of null-subject languages (see Guasti (2002: 137–9) for a summary); for Rizzi, this is because in null-subject languages infinitives must raise out of VP (see

Kayne (1991) and note 10 of Chapter 1 on movement of non-finite verbs), and this is of course impossible if the structure above VP is not there. Similarly, both analyses can account for the fact that clauses containing root infinitives cannot be V2 clauses (see (4b) above): in both cases this is because the C-position is not available as a target for movement; for Wexler, because non-finite verbs cannot raise there, and root infinitives are not specified as finite; for Rizzi, because CP is simply absent. The incompatibility of root infinitives with clitic and weak-pronoun subjects is again straightforward for Rizzi: it is well-known that these elements are attracted to the position bearing agreement features (which we take to be T), and so if this position is simply absent, such elements cannot be licensed. This restriction is problematic for Wexler's analysis, however. Property (4d), that root infinitives are not found in wh-questions, is once more straightforward for Rizzi: such questions clearly depend on the nature of the C-position, a position which is absent in a truncated clause. Wexler cannot handle this fact straightforwardly, since wh-questions may be either infinitival or finite, although of course infinitival wh-questions must be indirect questions in adult language: cf. *I don't know what to do/what I should do*, and in fact Wexler takes issue with this generalization about root infinitives, citing examples like *Where train go?* from child English. Intriguingly, though, such examples seem to be restricted to English.² Finally, both analyses can account for (4e) as long as it is assumed that auxiliaries require a fully-specified T-position, something lacking in the immature grammar on both analyses.

Root infinitives are typically no longer found after age three (Guasti 2002: 146). Both Wexler and Rizzi propose that this is due to the maturation of the grammatical system. They thus take the intermediate stages of language acquisition, or at least some of them, to represent literally immature grammars: grammars of a type that do not underlie any form of (non-pathological) adult linguistic behaviour. The idea that linguistic competence matures during the intermediate stages of language acquisition is, in the context of the assumption of a genetically-determined language faculty, quite reasonable. As Guasti says (2002: 146):

² There has been a debate about the analysis of such productions in child English. Roeper and Rohrbacher (2000) give numerous examples of the type *Where __ go?* Guasti (2002: 139, 202–8) provides a summary and references.

Maturation is likely to control some aspects of language development – for example, the fact that infants start to babble orally or manually around 6–8 months. According to the maturational view, a genetic program also controls the development of syntax . . . and determines the timing by which components of UG become available to the child. Under this view, R[oot] I[nfinitive]s occur because principles of UG have not matured.

The principles of UG in question clearly concern the fact that a main-clause Tense has to be fully specified for finiteness and other features, something both Rizzi and Wexler articulate in differing ways.³ So the account of root infinitives relies in one way or another on the idea that the requirement for specification of these features matures, typically, at around three years old.

In terms of these analyses, then, the root-infinitive phenomenon does not come close to our ‘ideal scenario’ for relating syntactic change and the acquisition of syntax. It simply involves the transition from one stage of acquisition to the next owing to the genetically-determined maturation of UG principles determining the well-formedness of main-clause Tense. However, as mentioned above, another way of moving from one stage of acquisition to the next must be the incorporation of further data: the accumulation of experience through greater exposure to the PLD. This idea and the maturation approach are not incompatible: in fact, it is natural to think that there is positive feedback between continued exposure to PLD and maturation of aspects of UG, in that greater exposure to data may in fact cause the system to mature as long as a certain age threshold has been passed. This view is supported by the fact that it is known that there is a **critical period** for language acquisition in general (see §5.4 for discussion of recent evidence for this), and that environmental stimulus is required in order for the system to come into operation at all.

In these terms, we might conceivably relate the root-infinitive phenomenon to our ideal scenario by adopting a two-stage approach to the acquisition of verbal agreement. Suppose that the ‘first pass’ acquirers make to the acquisition of agreement, at a rather early stage of acquisition, involves setting the null-subject parameter. Thus, if the relevant kind of agreement

³ If the minimalist conception of UG is taken on board, there may be rather few components of UG which are in principle available to come ‘on line’ at different stages of language acquisition, something which would impose inherent restrictions on how far maturation could be invoked to explain properties of child production. It is not clear how far this affects the analyses of root infinitives under consideration here, though.

and other properties (see §1.2.1) are expressed in the PLD, the positive value of the null-subject parameter is expressed and thus set by the acquirer. At this point, let us suppose that the immature system treats all non-null-subject systems alike: as having no agreement. Now, in §2.1, (16), I suggested, following Vikner (1997), that a certain pattern of verbal agreement, less robust than that required for a positive value of the null-subject parameter and yet greater than zero, expresses a positive value for the V-to-T parameter. At this intermediate stage, where all non-null-subject languages are treated as entirely lacking in agreement, there can therefore be no verb-movement. This can provide an account of the root-infinitive stage. Recall the basic properties of root infinitives, given in (4) above, repeated here:

- (4)
- a. Root infinitives do not occur in null-subject languages.
 - b. Root infinitives are not introduced by nonsubject XPs in V2 languages.
 - c. Root infinitives are incompatible with clitic and weak-pronoun subjects.
 - d. Root infinitives occur in declaratives, but not in wh-questions.
 - e. Root infinitives are incompatible with auxiliaries.

Clearly, on this view (4a) is accounted for, as a central assumption is that root infinitives result from an immature negative setting of the null-subject parameter. To the extent that clitics and weak pronouns depend on the presence of strong agreement, then (4c) is accounted for. Properties (4b) and (4d) depend on verb-movement to C (recall that root wh-questions involve verb- or auxiliary-movement to C in all the languages in question), and this will not be available if V-to-T movement is not available (and if auxiliaries are not merged directly in T, see below). Finally, we can account for (4e) if we assume that auxiliaries are elements which must always either move to T or be merged there (see §2.1), independently of the parameter determining V-movement to T. So the root-infinitive phenomenon could conceivably arise if, in acquiring systems with *little* verbal inflection, children assume there is *none at all* at first. This idea is similar in some respects to Phillips' (1995) idea that root infinitives arise from the failure of the features of V and T to combine either through Move or Agree. The phenomenon disappears when children make a later, 'second pass' at the acquisition of agreement, and at this stage they are sensitive to the expression of the agreement present in some non-null-subject languages. (This second pass may arise either through exposure or maturation, or, most likely, a combination of the two as described above.) This second pass is

related to the acquisition of the parameters involving V-movement, particularly V-to-T movement. The idea that verbal agreement is acquired in two stages is supported by the fact that there is good evidence that children acquire the agreement marking in null-subject languages early (Guasti (2002: 120–2) and the references given there), while it is well known that the much-impooverished agreement marking in English is acquired much later (Cazden 1968, Brown 1970); indeed, there is some reason to think that English agreement morphology is acquired like irregular tense marking (Maria-Teresa Guasti (p.c.)). The situation in German, on the other hand, appears equivocal, with Wexler and Poeppel (1993) arguing for early acquisition of agreement marking and Clahsen and Smolka (1985) and Clahsen and Penke (1992) arguing that there are many errors in the early use of German agreement marking (Guasti *loc cit.*).

The interest of the account of root infinitives just sketched is that it can get us close to our ideal scenario for linking acquisition and change. In terms of the scenario, we can take L to be German and L' to be English. English is syntactically innovative in relation to German, in that it is probable that the common parent language, Proto-West Germanic, set the V-to-T parameter to the movement value, as we have assumed for German (see §1.3.1, notes 18 and 30). The reason for this is that Proto-West Germanic is usually reconstructed as having a very rich verbal agreement system, and hence enough agreement was present to express a positive value for the V-to-T parameter. Hogg (1992: 147ff.) presents a reconstructed stage of pre-OE; and the paradigms of the Gothic verb given in Jasanoff (2004: 900) nearly all have forms of both strong and weak verbs which distinguish all person-number combinations. Modern English, of course, sets this parameter to the negative value. Second, we observe that early production in German shows a tendency for strings which appear to express value v_j for P: these are the root infinitives.

So far, we are close to making a connection between the two areas of acquisition and change. However, the third step of our scenario involves the observation that the 'apparently aberrant production in L', i.e. the root infinitives, ceases when some further feature F of L is acquired. If we could establish that German agreement marking is acquired relatively late, and coincides with the loss of root infinitives, this would be just what we need. However, as mentioned above, there is a debate regarding the timing of the acquisition of German agreement (possibly because the researchers in question made use of differing experimental methodologies; Maria-Teresa

Guasti (p.c.)). And so here we cannot be sure of attaining the ideal scenario.

The fourth part of the scenario is straightforward, on the other hand. We have already observed the correlation between the loss of agreement marking (F) from English (L') when the V-to-T parameter changed value.

We see that it may be possible to link root infinitives to some aspects of the loss of V-to-T movement in English, although making the connection is not without problems.⁴ Let us now turn to the other L1-acquisition phenomenon of interest: early null subjects.

Early null subjects are illustrated in (5) (examples from Guasti (2002: 151); sources are given there):

- (5) a. Se, blomster har. (Swedish)
 see, flowers have/has
 'Look, (I/you/she/we) have/has flowers.'
 b. Tickles me.
 c. Mange du pain. (French)
 eat-3sg some bread

While similar to root infinitives, early null subjects differ from them in that the verb is clearly finite, as can be seen from the forms in (5), and the fact that they are compatible with the presence of auxiliaries (compare (4e) above).

At first sight, it might seem that early null subjects provide evidence that the null-subject parameter is not set as early as we have been supposing up to now, essentially following Guasti. In fact, in her pioneering work on this

⁴ Arguably the biggest problem is that Guasti (2002: 166) shows that there is an implicational relation between root infinitives and early null subjects: if a child's grammar allows root infinitives, it allows early null subjects. But if root infinitives result from an early, correct, negative setting of the null-subject parameter, then this seems paradoxical. One way to handle this would be to claim that the correct acquisition of the strong agreement morphology associated with a positive value of the null-subject parameter implies that the null subject can be syntactically represented in such systems (either as *pro* or in the verbal inflection itself – see §1.1.1), while the immature negative setting associated with complete lack of agreement implies that overt subjects are simply not subject to a licensing condition, while covert subjects must be inferred from argument structure of the verb (i.e. as a kind of 'implicit argument'). It seems that any analysis of root infinitives has to allow for the fact that subjects, when structurally manifested, cannot be licensed in the way they are in the adult system (for example, by Agree with T's φ -features as discussed in §2.3). In different ways, this is true for Wexler's and Rizzi's analysis of root infinitives, as well as the one sketched above.

phenomenon, Hyams (1986) proposed that early null subjects were an indication of ‘parameter-missetting’ in relation to the null-subject parameter, in that children acquiring non-null-subject languages initially set the parameter to the positive value. This led to the suggestion that the null-subject parameter may have a default positive value, a matter we return to briefly in §3.4 below.

Hyams’ account of early null subjects comes close to fulfilling our ideal scenario for the connection between language acquisition and language change. In terms of that scenario, we could suppose that *L* is Italian and *L'* is English; *P* is the null-subject parameter; v_i is the positive value of that parameter; and v_j is the negative value. Since it is likely that Proto-Indo-European, the common parent of English and Italian, was a null-subject language (see §4.4.4), then English is syntactically innovative with respect to Italian as far as this parameter is concerned. The crucial factor *F*, causing children to fix the parameter correctly after a period of ‘aberrant’ production, could be either the presence of modal auxiliaries (originally identified by Hyams as inherently incompatible with null subjects) or overt expletive pronouns (often thought to be incompatible with a positive value for the null-subject parameter – see Rizzi (1986a), and, for a different view, Holmberg (2005)). The difficulty with this is that feature *F*, which, following Hyams, we take to be expletive pronouns or modal auxiliaries, is predicted to have arisen when English, or the relevant ancestor of English, ceased to be a null-subject language. But, first, we have little clear idea as to when that was; as just mentioned, it is likely that Proto-Indo-European was a null-subject language, and it is possible that Proto-Germanic was (see the above comment on archaic Germanic verbal agreement). As for the likely status of the null-subject parameter in the runic inscriptions (which may represent an early form of either North or Northwest Germanic, see Faarlund (2004a: 908)), Faarlund (2004a: 920) argues that the data from the surviving runic inscriptions is too sparse for any conclusions to be drawn. Second, both modal auxiliaries and overt expletive pronouns are innovations in the recorded history of English. It is usually thought that modal auxiliaries of the Modern English type arose in the sixteenth century (as was briefly discussed in §2.1; see note 4 of that section on the chronology of this change), while overt expletives appear during ME (A. Williams 2000; Biberauer 2003; Biberauer and Roberts 2005a). Thus there is a clear chronological mismatch regarding this feature, and so this vitiates this particular application of the scenario.

An alternative might be to take L' to be French. Here we are on much firmer ground regarding the parent language: it is clear that Latin was a null-subject language, and so French can be defined as syntactically innovative in relation to Italian in this respect. It is also clear that French shows early null subjects (see (5c)), and so we see the relevant kind of ‘aberrant’ production in child language. However, we run into difficulty with the third part of the scenario: neither French nor Italian has modal auxiliaries, and so the relevant factor must be overt expletive pronouns, which of course Modern French has but Italian lacks. But the problem is that OF, which, as we saw in §1.2.2, was a null-subject language, had overt expletives, as the following examples show:

- (6) a. Il est juget que nus les ocirum.
it is judged that we them will-kill
(*Roland*, 884; Roberts 1993a: 150)
- b. Il ne me chaut.
it not to-me matters
(Einhorn 1974: 123)

One possibility for saving this approach might be to claim that the expletives illustrated in (6) are in SpecCP (which they almost certainly are, given the V2 nature of OF; see §1.3.2), and that the relevant property for changing the null-subject parameter, in both acquisition and change, involves expletives in SpecTP.

However, there are examples of expletives in SpecTP in OF, such as the following:

- (7) car ainsin estoit il ordonne
for thus was it ordained
‘for thus it was ordained’
(Vance 1988 (26b), 159; Roberts 1993a (102b): 147)

We must therefore conclude that OF had expletive pronouns. We therefore do not know what caused the null-subject parameter to change its value in the history of French, and so we are unable to relate this change to the acquisition of a given value of the null-subject parameter.

The fundamental problem with the Italian–French comparison is that, since Hyams’ early work, evidence has emerged that early null subjects are not the result of a ‘missetting’ of the null-subject parameter to the Italian value. The main reason for this is that early null subjects do not occur in the following environments:

- (8) a. questions with a fronted wh-element;
 b. subordinate clauses;
 c. matrix clauses with a fronted non-subject.

On the other hand, null subjects readily occur in these environments in adult null-subject languages, as the following Italian examples (from Guasti (2002: 159)) show:

- (9) a. Cosa __ hai detto? (wh-question)
 what have-2sg said
 'What did you say?'
 b. Gianni ha detto che __ verrà. (subordinate clause)
 John has said that (he) will come.
 c. Ieri __ ho parlato a Carlo. (root clause with fronted adverb)
 Yesterday (I) have spoken to Carlo.

Because of data like this, it has been widely concluded that, despite initial appearances, early null subjects are not a case of missetting of the null-subject parameter to the 'Italian' value. Hence our comparison of Italian and French above in relation to the 'ideal scenario' was to no avail.

Another option, pursued by Hyams (1992), was to claim that early null subjects result not from a 'subject-drop' option of the familiar Italian kind, but from a 'topic-drop' option of the kind seen in languages such as Chinese and Japanese (see in particular Huang (1984; 1989) on this). The advantage of this idea is that it reconciles the occurrence of early null subjects in languages with impoverished agreement systems with the known facts of adult languages: while null-subject languages like Italian appear to require 'rich' verbal agreement for the recovery of the content of null subjects (see §1.1.1), topic-drop languages like Chinese and Japanese have no agreement at all and yet allow null arguments of various kinds, as the following Chinese examples illustrate:

- (10) a. __ kanjian ta le.
 (he) see he Asp
 b. Ta kanjian __ le.
 he see (him) Asp
 'He saw him.'

The disadvantage of this approach is also apparent from (10b): topic-drop languages allow null objects fairly freely, in addition to null subjects. On the other hand, Hyams and Wexler (1993) show that early null objects are rather rare in child English, and Wang *et al.* (1992) show that child Chinese

allows null objects significantly more freely than child English. (For further discussion and statistical evidence, see Guasti (2002: 157–8).) So the idea that the putative parameter-missetting is in the ‘Chinese’ direction rather than in the ‘Italian’ one does not appear to hold up either.

So it appears that there is no obvious ‘parameter-missetting’ going on with early null subjects. Other possibilities which have been explored to account for this phenomenon include relating it to the **diary drop** we briefly saw in §1.1.1. The relevant examples are repeated here:

- (11) a. ... cried yesterday morning.
(Plath 1983: 288)
b. Elle est alsacienne. ... Paraît intelligente.
She is Alsatian. Seems intelligent.
(Léautaud 1989: 48)

These examples have been argued to involve clausal truncation (but at a higher level of clausal structure than that involved in root infinitives as described above) by Haegeman (2000) and Rizzi (2000). Indeed, a truncation analysis seems to account for this phenomenon quite well; see Guasti (2002: 166ff.) for summary and discussion. The other accounts which have been put forward involve extra-syntactic factors, such as processing difficulties (Bloom 1990) and metrical difficulties (dropping of weakly stressed syllables) (Gerken 1991). As these do not involve factors which may be relatable in any direct way to parametric change, I will leave them aside here. (Once again, see Guasti (2002: 179–83) for discussion.)

In conclusion, for all their intrinsic interest and the light they shed on L1 acquisition, it seems that neither root infinitives nor early null subjects can clearly be related to the kinds of phenomena known in parametric syntactic change. Hence no clear connection can be made between studies of immature competence and the acquisition-driven conception of parametric change.⁵ Although perhaps disappointing, this is not surprising, and does

⁵ Rizzi (2005: 97–100) conjectures that root infinitives and early null subjects (as well as ‘determiner drop’ and ‘copular drop’, two other features of the production of two- and three-year-olds not discussed here) may arise from the fact that ‘the child initially assumes all the parametric values which facilitate the task of the immature production by reducing computational load’ (97). This general strategy is constrained by the values of the parameters which are fixed early, hence the observation that the null-subject parameter is correctly fixed in the acquisition of null-subject languages, but early null subjects and root infinitives may appear in the acquisition of non-null-subject languages owing to the adoption of the conjectured strategy.

not, at least in my view, prove the acquisition-driven conception of parametric change wrong (although very interesting potential evidence that it is right is thereby sadly lacking).

There are various explanations for this state of affairs. First, no-one has really looked carefully for a connection between acquisition and change: there is something of a sociological divide between linguists working on L1 acquisition and those working on diachronic syntax.⁶ This is a regrettable, but entirely contingent state of affairs, and something which can in principle easily be remedied. Second, good empirical coverage of early production is limited to a few languages: English, French, Dutch, German, and Italian most prominent among them to judge by Guasti (2002); in diachronic syntax, selected topics have been studied in the histories of a range of languages, but a good overall picture of the syntactic history of very few languages other than English and French is hardly available. Thus our database of languages is at present extremely small, and so our chances of finding the ideal case correspondingly restricted. Third, and most importantly, the nature of the data in both cases may make the ideal scenario described above hard to identify. The immature competence of small children goes hand in hand with a general cognitive immaturity, notably for example a smaller short-term memory capacity, which means that comparing children's grammars with adult grammars may really be like comparing chalk with cheese. The diachronic data we have is the output of adult competence, but of course the surviving texts have been subject to many vicissitudes of history; one of the principal goals of traditional philology is simply to unravel the sometimes tortuous histories of extant texts. And so in their different ways both the acquisition data and the diachronic data are corrupt, and this makes comparing data from the two sources in any reliable way all the more difficult. Of course, what we would ideally like is an acquisition study of an earlier stage of a language. Since

⁶ As mentioned several times already, Lightfoot has consistently made the connection between syntactic change and the acquisition of syntax. In particular, Lightfoot (1991) develops the 'degree-0 learnability' theory with a view to explaining aspects of both. Lightfoot's application of degree-0 learnability to word-order change in the history of English was briefly discussed in §2.5; he also applies the same notion to aspects of language acquisition. This, Clark and Roberts (1993), and Roberts (1999) are the only cases in the literature where an explicit connection is attempted between change and acquisition (although see also DeGraff (1999)), although the connection is mentioned in Hyams (1986: 23, n. 1).

acquisition studies only began in the mid-twentieth century (see Grégoire (1937–47); Jakobson (1941); Leopold (1939–49)), this is not possible.⁷

Although the conclusion of this section may seem pessimistic, three points are worth bearing in mind. First, as stated above, what is lacking is empirical confirmation through acquisition studies that parameter change is driven by acquisition; the lack of evidence does not disconfirm this approach, especially since we can quite easily see why such evidence may be lacking, and the conceptual arguments (see §2.1) are unaffected. Second, the ‘ideal scenario’ may well be realized by some new data; in fact, the root-infinitive evidence comes close to this, as we saw above, and an enhanced understanding either of this data or of parameter changes involving agreement may yet yield that scenario. Third, as mentioned above, Guasti’s conclusions regarding early and accurate setting of several important parameters, listed in (2), lead us to pose the intriguing logical problem of language change. It is to this last issue that we now turn.

3.2. The logical problem of language change

In this section, we will look in more detail at the apparent tension we noted in the previous section between the evidence that many important parameters are set at an early stage of language acquisition and the idea that syntactic change is driven by abductive reanalysis, associated with parametric change, of parts of the PLD. This will lead us to formulate the logical problem of language change, as a kind of paradox for learnability theory. (At the same time, some basic concepts of learnability theory are introduced.) In considering a possible approach to solving this problem, we

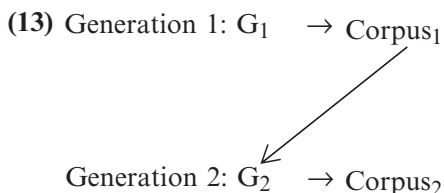
⁷ I am aware of one striking exception to this generalization: the journal kept by Jean Héroard in the period 1601–28 of the speech of the young dauphin Louis XIII (Ayres-Bennett 1996: 216ff.) Héroard, who was the dauphin’s personal physician, transcribed samples of the dauphin’s speech between the ages of 3;3 and 9;3 (i.e. in the years 1605–10). This immediately yields some interesting observations, notably that *ne* was already dropped, i.e. the dauphin’s production represented a very early instance of Stage III of Jespersen’s Cycle (Ayres-Bennett 1996: 221; see §2.2, note 8 and §1.4.2). Ernst (1985) is an edition of Héroard’s journal. Ayres-Bennett (2004: 185ff.) provides further examples of the sporadic omission of *ne* in seventeenth-century French.

introduce the **Inertia Principle** first put forward in early version of Keenan (2002), as developed in Longobardi (2001). This in turn leads to an explicit characterization of the conditions under which abductive change can take place, in terms of ambiguity and complexity. The discussion will sharpen some of the notions connected to parameter-setting that we are concerned with, and so our overall account of syntactic change will be further elucidated, and a number of questions raised for detailed consideration in later sections of this chapter.

Let us begin by recapitulating some of the ideas we have put forward up to now:

- (12) a. The central mechanism of syntactic change is parameter change.
 b. Parameter change is manifested as (clusters of) reanalyses.
 c. Reanalysis is due to the abductive nature of acquisition.

Concept (12a) was argued for at length in Chapters 1 and 2, and we noted in §2.3 that parameter change may not be the only mechanism of syntactic change, although it does appear to be the principal one. Both (12b) and (12c) were introduced in the discussion of reanalysis in §2.1. There we also saw Andersen's (1973) familiar schematization of abductive change, which I repeat here:



Owing to the abductive nature of language acquisition, G_2 may in principle not be identical to G_1 . If identity between grammars is defined in terms of identity of parameter-settings, this implies that G_2 may differ from G_1 in at least one parameter value, and, by (12b), this will give rise to a cluster of reanalyses. Following Roberts and Roussou (2003: 11), we can give the following characterization of abductive change (cf. also Lightfoot (1979; 1991; 1999)):

- (14) (A population of) language acquirers converge on a grammatical system which differs in at least one parameter value from the system internalized by the speakers whose linguistic behaviour provides the input to those acquirers.

This essentially states what the schema in (13) illustrates. So where G_2 may differ from G_1 in at least one parameter value in (13), an abductive change takes place.

Let us consider (13) and (14) in the light of the main concepts of learnability theory, the abstract, formal theory which deals ‘with idealized “learning procedures” for acquiring grammars on the basis of exposure to evidence about languages’ (Pullum 2003: 434). In terms of this theory, any learning situation can be characterized in terms of the answers to the following questions (this presentation is from Bertolo (2001: 2ff.)):

- (15) a. ‘What is being learned . . . ?
b. What kind of hypotheses is the learner capable of entertaining?
c. How are the data from the target language presented to the learner?
d. What are the conditions that govern how the learner updates her responses to the data?
e. Under what conditions, exactly, do we say that a learner has been successful in the language learning task?’

The answers to some of these questions are obvious, given the assumptions we are making here regarding UG and parametric variation: for example, the answer to (15b) is that the learner can only consider distinct values of parameters in the acquisition of syntax. Similarly, the answer to the question of how the data are presented to the learner – (15c) – appears to be simply in the form of spontaneous linguistic behaviour which makes up the PLD. As we saw in the Introduction to Chapter 1, no negative evidence is available to the learner. One could put the answer to this question more abstractly, and say that the data is presented in the form of P-expression – as defined in (8) of §2.1 – in the strings in the PLD.

Some of the other questions in (15) are trickier. For instance, our entire discussion of the possibility of evidence of ‘parameter-missetting’ in the previous section can be construed as addressing (15d). On the basis of that discussion, we have to conclude that no updating of parameter values takes place, which might be relevant. But it is (15a) and (15e) which are most important to our concern with the relation between learnability/acquisition and change. One possible answer to (15a) is that acquirers are learning the parameter values of the grammar that produces the PLD. But in that case the learning task would be seen as unsuccessful when abductive parametric change of the sort schematized in (13) takes place. This seems to be the wrong conclusion, since the learners in (13) have acquired *a* grammar, just not the parental one (since $G_1 \neq G_2$). So the answer to (15a) would be simply that a parametric system is learned. This point relates to (15e), too: the criterion for successful learning cannot be replication of the parental grammar, but approximation to it, in such a way as abductive change of the

sort shown in (15) is possible. For this last reason, it is often thought that **Probably Approximately Correct (PAC) algorithms** are **learning algorithms** which can provide useful simulations of language acquisition. (See Bertolo (2001: 8–10); Clark and Roberts (1993); Pullum (2003: 433); and in particular Niyogi (2004: 75ff.) and the references given there.)

Now, most work on L1 acquisition assumes that the stable state of acquisition, S_S , corresponds exactly to the target grammar. In other words, it assumes that no ‘mismatch’ arises between G_1 and G_2 in (13). In considering the stages of L1 acquisition in the previous section, we implicitly took this view in saying ‘we can consider that the various intermediate states differ from one another in representing successively closer approximations to the adult system (S_S) in terms of the values of the parameters. To put it another way: if m parameters are set to the adult value at stage S_n then at least $m + 1$ parameters are set to the adult value at stage S_{n+1} .’ But, as we have seen, abductive change requires a slightly looser formulation than this, in order to allow for the possibility of ‘mismatch’ between G_1 and G_2 in (13).

The following remark by Niyogi and Berwick (1995: 1) summarizes the difference between the standard assumptions in work on language acquisition and what seems to be required for an acquisition-driven account of change: ‘it is generally assumed that children acquire their ... target ... grammars without error. However, if this were always true, ... grammatical changes within a population would seemingly never occur, since generation after generation children would have successfully acquired the grammar of their parents’.

Thus language acquisition is usually taken to be deterministic in that its final state converges with the target grammar that acquirers are exposed to. The postulation of abductive change challenges exactly this assumption.

In the previous section, we saw good support for the deterministic assumption in L1 acquisition. Recall Guasti’s list of parameter values which appear to be correctly fixed from roughly the time of the earliest multiword utterances:

- (2) a. the value of the head direction parameter;
- b. the value of the V-to-T parameter;
- c. the value of the topic-drop and null-subject parameters;
- d. the value of the parameters governing question formation, the one governing overt movement or *in-situ* placement of the wh-element and the one regulating T-to-C movement (inversion).

Moreover, we saw in the previous section that two of the best studied phenomena of child production, root infinitives and early null subjects, are not easily or obviously analysed in terms of ‘parameter-missetting’, i.e. in terms of mismatches between G_1 and G_2 in (13). As Roberts and Roussou (2003: 12) put it: ‘the standard paradigm for language acquisition is not immediately compatible with the observation that grammatical systems change over time.’ This ‘standard paradigm for language acquisition’ is more than just a methodological simplification on the part of linguists working on L1 acquisition: (2) gives the evidence of accuracy and earliness in the setting of a number of important parameters. And yet we saw in Chapter 1 that all the parameters referred to in (2), with the exception of topic drop, have changed their values in the recorded histories of various languages. So we see a tension between the results of L1 acquisition research and what we can observe about syntactic change.

The obvious explanation for the fact that children are able to set many parameters very early lies in the highly restricted range of analyses of the PLD that UG allows them to entertain and the limited exposure to PLD needed for parameter fixation. (In this respect, the facts reported in (2) support the argument from the poverty of the stimulus, as noted in the previous section.) But then, how are we to explain the fact that these parameters are subject to change over time? Notice that our answers to the learnability questions in (15) do not answer this question; they simply allow for abductive parametric change without explaining how it happens.

This leads us to the logical problem of language change, which we can formulate as follows (this formulation is based on unpublished work with Robin Clark (Clark and Roberts 1994: 12)):

- (16) If the trigger experience of one generation permits members of that generation to set parameter p_k to value v_i , why is the trigger experience produced by that generation insufficient to cause the next generation to set p_k to v_i ?

The logical problem of language change as formulated here is close to the Regress Problem for reanalytical approaches to change which we introduced in §2.1 in the following terms: ‘an innovation in Corpus_2 [in (13)] may be ascribable to a mismatch in G_2 (compared to G_1), but it must have been triggered by something in Corpus_1 – otherwise where did it come from? But if Corpus_1 could trigger this, then how could G_1 produce this property without itself having the innovative property?’ Essentially, this

formulation puts the problem the other way around as compared to (16), in addition to not being directly phrased in terms of parameter change. I take it, though, that (16) subsumes what I called the Regress Problem in the earlier discussion.

The first thing that is required in order to find our way out of the dilemma stated in (16) is a slightly weaker notion of the deterministic nature of L1 acquisition than that which is usually assumed in the L1-acquisition literature. So let us propose, following Roberts and Roussou (2003: 13), that ‘the goal of language acquisition is to fix parameter values on the basis of experience; all parameter values must be fixed, but there is no requirement for convergence with the adult grammar’. More precisely, as mentioned in the discussion of (15e) above, let us suppose that the goal of acquisition is to *approximate* the parental grammar, not to *replicate* it. Making this move allows for p_k in (15) to receive a different value from that found in the input, therefore making space for language change. The stable state S_S of language acquisition now amounts to the situation where all parameters are fixed to a given value (cf. the remark in relation to parameter-setting in §1.1 that ‘not deciding is not an option’). Let us call this view of the endpoint of language acquisition ‘weak determinism’.

The ‘approximation’ approach may seem too weak, in particular in that it does not appear to account for the results of L1-acquisition research as summarized in (2), since it in principle allows parameters to vary freely and randomly from generation to generation. However, Roberts and Roussou (2003: 13) add an important proviso to the above quotation to the effect that convergence with the adult grammar ‘happens most of the time’; that is, approximation usually amounts to replication. This brings us to an important principle of syntactic change, first put forward in Keenan (2002): the Inertia Principle. Keenan formulates this as follows:

- (17) Things stay as they are unless acted on by an outside force or decay.
(Keenan (2002:2))

This principle is very general; in fact it holds of the physical world in general, taking decay to include entropy, i.e. the second law of thermodynamics. For our purposes, we can take (17) to mean that, although L1 acquisition is not inherently deterministic but rather weakly deterministic in Roberts and Roussou’s sense, the target system is successfully converged on, i.e. the stable state S_S of acquisition has the same parameter values as that of the parent system; G_1 and G_2 in (13) do not differ. This is no doubt

due to the highly restricted range of analyses of the PLD that UG allows and the limited exposure to PLD needed for parameter fixation, i.e. standard poverty-of-stimulus considerations.

Longobardi (2001: 278) adopts Keenan's principle, and puts forward the following very interesting version of it:

- (18) 'syntactic change should not arise, unless it can be shown to be *caused*'
(emphasis his)

In other words, as Longobardi says, '*syntax*, by itself, is diachronically completely inert' (277–8). It is clear that this view is compatible with the results of L1-acquisition research, as reported in the previous section and in (2). If we combine (18) with Roberts and Roussou's weak determinism, we arrive at the following:

- (19) If a definite value v_i is expressed for a parameter p_i in the PLD, then
(a population of) acquirers will converge on v_i .

In other words, given adequate P-expression, inertia will hold. So inertia implies that most of the time abductive change does not happen. P-expression was introduced and defined in §2.1; the definition, and the corollary definition of trigger, is repeated here:

- (20) a. Parameter expression:
A substring of the input text S expresses a parameter p_i just in case a grammar must have p_i set to a definite value in order to assign a well-formed representation to S.
b. Trigger:
A substring of the input text S is a trigger for parameter p_i if S expresses p_i .

As long as there is a trigger for a given parameter value, then, inertia will hold and abductive change will not take place.

Under what circumstances does abductive change happen, then? This must be when no definite value v_i is expressed for a parameter p_i in the PLD. According to Longobardi's version of Inertia in (18), this lack of robust P-expression must be 'a well-motivated consequence of other types of change (phonological changes and semantic changes, including the appearance/disappearance of whole lexical items) or, recursively, of other syntactic changes' (278).

More precisely, we propose, following our discussion of reanalysis in §2.1, that both ambiguity and opacity of the P-expression are required in

order for abductive change to take place. Ambiguity is defined in relation to parametric systems as follows:

(21) a. P-ambiguity:

A substring of the input text S is strongly P-ambiguous with respect to a parameter p_i just in case a grammar can have p_i set to either value and assign a well-formed representation to S .

b. A strongly P-ambiguous string may express either value of p_i and therefore trigger either value of p_i .

c. A weakly P-ambiguous string expresses neither value of p_i and therefore triggers neither value of p_i .

In fact, as we saw in §2.1, strong P-ambiguity is what is required for reanalysis. These definitions are repeated from that discussion, where they were illustrated in relation to certain reanalytical changes. Weak P-ambiguity arises where some parameter value is undetermined, also possibly leading to change in a parameter value, although not necessarily through reanalysis.

We can define opacity in terms of complexity (see Lightfoot (1979)). Following an idea developed in Clark and Roberts (1993), let us assume that learners are conservative in that they have a built-in preference for relatively simple representations (the precise characterization of simplicity will be discussed in §3.4 and §3.5). If a given piece of PLD is P-ambiguous, there will be at least two representations for it, each corresponding to a different grammar, i.e. representing systems with distinct parameter values. Assuming that any two representations differ in complexity and therefore opacity, the learner will choose the option that yields the simpler representation. The more complex representation will be both opaque (in virtue of being more complex than the other available representation(s)) and ambiguous (by definition). Therefore it is inaccessible to the learner, i.e. it is effectively unlearnable. The Inertia Principle tells us that the strong P-ambiguity of the trigger (and therefore the relative opacity, assuming that any two alternative representations differ in overall complexity) arises through either extra-syntactic factors or as the consequence of an independent syntactic change.

Actually, closer reflection reveals that the circumstances just described do not guarantee a change; they merely suspend inertia, since we can take it that P-expression (and therefore strong P-ambiguity) and the preference for relative simplicity are forces acting on the learner, in the sense relevant for the Inertia Principle as stated in (17). Hence, it is possible that things

will not stay as they are. Whether a parameter value of G_1 actually changes will depend on the relative complexity of the representations of aspects of the PLD entailed by the parameter-settings in G_1 compared to those of G_2 : a parameter will change if it corresponds to the single option expressed by the PLD for G_1 and the more opaque of two options expressed by the strongly ambiguous PLD for G_2 . Here again weak determinism is relevant, in that it implies that under these conditions a definite value v_i for p_i will be assigned. This value will still be compatible with the input, but – again thanks to weak determinism – may differ from that of the target grammar, in which case an abductive change takes place. So we see that the simplicity metric is the ‘safety mechanism’ alluded to earlier.

So, our tentative answer to the question posed in (16) is that between the two generations in question there is a change in the PLD. In other words, some extra-syntactic factor, or at least a factor independent of the change in question, introduces P-ambiguity into the expression of at least one parameter in the PLD of G_2 . Still assuming that any two representations can be distinguished in terms of complexity (and still leaving complexity undefined, for the time being), complexity/opacity will then choose between the two values, possibly leading to a parametric change.⁸ The crucial question becomes that of locating the factors that may introduce P-ambiguity into the PLD. This is what we turn to in the next section.

Clearly, all of the above discussion turns on the notion of complexity, which we must therefore now define. The commonest way of determining this is by simply counting some aspect of a derivation or representation. (See in particular Chomsky and Halle (1968, Chapter 9) on complexity and markedness in phonological systems; we return to the discussion of markedness in §3.4.) As Roberts and Roussou (2003: 200) point out, syntactic representations offer several possibilities:

In principle, there are several formal options available in syntactic representations or derivations: one could count nodes, branching nodes, traces [i.e. copies – IGR], chain links, symbols or features.

⁸ Kroch (2000: 700) points out that it is also possible that ‘extrasyntactic’ change may be attributable to some property of the learner, for example, age at the time of acquisition. This is relevant ‘in the case of change induced through second-language acquisition by adults in situations of language contact’ (*ibid.*). I will discuss this case briefly in the next section, returning to it in detail in Chapter 5.

After considering the various options, Roberts and Roussou opt for a feature-counting approach; this is in fact very much in the original spirit of Chomsky and Halle (1968). Here I give a slightly simplified version of Roberts and Roussou's proposal (for the original, see Roberts and Roussou (2003: 201)):

- (22) Given two structural representations R and R' for a substring of input text S, R is simpler than R' if R contains fewer formal features than R'.

The notion of 'formal feature' here is the standard one in current versions of syntactic theory, as introduced in §1.4.1: it includes features such as Person, Number, Gender, Case, and Negation. And, as we saw in §2.5, movement takes place where the Probe of an Agree relation has uninterpretable formal features and 'an extra property' triggering movement. Chomsky (2000; 2001) proposes that that 'extra property' is in fact a further formal feature known as the EPP feature. This means that Probes, in terms of formal features, are more complex than non-Probes, and Probes that cause movement are more complex than those which do not. We will see the effects of this definition of complexity in more detail in the next section.

In this section, we resumed certain aspects of the discussion of reanalysis in §2.1, notably the question of abductive change. Applying this idea strictly to parameter change, we arrived at the logical problem of language change as stated in (16), partly on the basis of some of the observations regarding language acquisition made in the previous section. We suggested a way of solving this problem on the basis of the Inertia Principle of (17) and the corollary stated by Longobardi in (18). We are led to the conclusion that abductive parametric change only occurs when the trigger for a given parameter value, as defined in (20b), is both ambiguous and opaque. P-ambiguity is defined in (21) and opacity/complexity in (22). P-ambiguity can only be introduced through extrasyntactic factors, for example, through language contact (see note 8), morphophonological erosion, or through an independent syntactic change.

In the remaining sections of this chapter, I will address the two main issues that this account of change raises: the nature of changes to the trigger, and the nature of complexity as the principal force which acts on parameter-setting, preventing things from staying the same (see (17)); this will be linked to the concept of markedness of parameter values. Finally I will draw these threads together and attempt a formal characterization of parameters.

3.3. The changing trigger

Following on from our rather abstract discussion of the logical problem of language change, this section looks at the idea that changes in the trigger experience – in particular the introduction of strong P-ambiguity – are responsible for language acquirers resetting parameter values. In other words, we investigate what both Kroch (2000: 700) and Longobardi (2001: 277–8) put forward implicitly as the solution to the problem: the idea that intergenerational changes in the PLD render earlier parameter-settings prone to abductive change. We saw that this is due to the expression of these parameters becoming ambiguous and opaque.

I will discuss three ways in which the PLD may be rendered ambiguous and/or opaque: contact-driven parameter-resetting, as suggested in King (2000), Kroch and A. Taylor (1997) and Kroch, Taylor, and Ringe (2000); cue-based resetting of the type advocated by Dresher (1999) and Lightfoot (1999) (the cue-based model was originally proposed by Dresher and Kaye (1990)); and morphology-driven parameter-resetting, as suggested in Roberts (1985; 1999); Roberts and Roussou (2003). These ways of rendering the PLD ambiguous and opaque are not mutually exclusive; it is very likely that all three possibilities exist.

3.3.1. *Contact-driven parameter-resetting*

The contact-driven view of parameter-resetting can be construed, in our terms, as the case where PLD is affected by an alien grammatical system. What this means is that Generation 2 in the schema for abductive change in (13) is subjected to a different kind of PLD from Generation 1 in that Generation 2 receives PLD that either directly or indirectly reflects a distinct grammatical system (i.e. set of parameter values) from that which underlay the PLD for Generation 1.

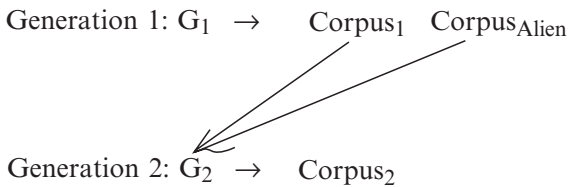
The direct case of contact is that where the PLD simply contains a significant quantity of tokens from a distinct system (where ‘distinct’ means that the grammar in question generates strings that cannot express the original grammar); this would naturally arise where Generation 2 is brought up in an environment which contains a language or dialect absent from the early experience of Generation 1. Such situations can and do arise through

many different types of historical contingency: emigration, invasion, and intermarriage being the most obvious but certainly not the only ones.

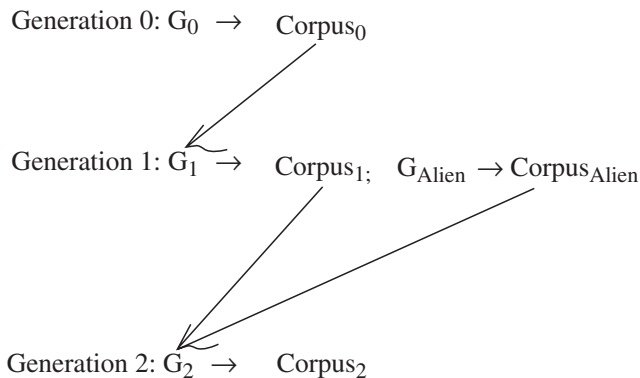
The indirect case of contact arises where Generation 1 uses a second language in interaction with Generation 2. Here the PLD for the two generations is very obviously distinct. In the case where this second language is a pidgin, Generation 2 may form a creole; I will leave this particular situation aside here and return to it in §5.3. If Generation 1 has learnt the second language after the critical period for language acquisition, the PLD will consist of **interlanguage** and many parameter-settings may be radically underdetermined by the PLD. (I will come back to this in the discussion of second-language acquisition and the nature of interlanguage in §5.1.) This situation gives rise to weak P-ambiguity and hence potentially to change, which may have the properties of the creation of new grammatical features *ex nihilo*.

The direct and indirect cases of language contact influencing PLD are diagrammed in (23):

(23) a. Direct contact:



b. Indirect contact:



As already mentioned, issues to do with the nature of language contact, second-language acquisition and possibly creolization arise here, but we will leave them aside and return to them in Chapter 5.

An example of what looks like direct contact, that is, contact-induced borrowing affecting aspects of syntax, comes from King's (2000) study of the French spoken on Prince Edward Island in Canada. Prince Edward Island (PEI) French is a variety of Acadian French, also spoken in the other Canadian Maritime Provinces (New Brunswick and Nova Scotia) and in parts of Newfoundland. French has been spoken in Acadia since the early seventeenth century (King 2000: 7), but according to the 1991 census, only 4.2 per cent of the population of PEI were native speakers of French, and only 2.3 per cent spoke the language at home. 86 per cent of these people lived in a single area, Prince County (King (2000: 19)). Contact with English is clearly very extensive, and King (2000, Chapter 6) documents much lexical borrowing and code-switching.

What is of interest in the present context is that King (Chapter 7) documents cases of Preposition-stranding in PEI French. As we saw in Box 1.5 of Chapter 1, Preposition-stranding is the cross-linguistically rather rare option of moving the complement of a preposition, while leaving the preposition 'stranded'. English and Mainland Scandinavian languages allow this, both in *wh*-questions, illustrated in (24a), and in passives, shown in (24b) (Icelandic in fact allows the equivalent of (24a), but not (24b) with a nominative subject: Kayne (1984: 117)):

- (24) a. Who did you speak to __?
b. John was spoken to __.

It is not clear what exactly permits this in English and the Scandinavian languages. In most other languages which have overt *wh*-movement, the preposition must move with the *wh*-phrase, i.e. it must be pied-piped (see §1.5 on the *wh*-movement parameter). This is the case of Standard French, for example (both examples in (25) are equivalent to (24a)):

- (25) a. *Qui as-tu parlé à __?
b. A qui as-tu parlé __?

Of course, English also allows the equivalent of (25b) (*To whom did you speak __?*). This is probably a case of formal optionality in English; a [+*wh*] C with an EPP feature can cause either the DP complement of the preposition or the whole PP to move, although Agree_{wh} holds just between

C and the DP. (We will look in more detail at the concept of formal optionality in §4.1.) Optionality has social value, in that the pied-piping option is characteristic of formal registers. French has some property which requires pied-piping of the DP in all cases. Kayne (1984) made an interesting and influential proposal concerning this, but in the context of technical assumptions which have not been carried over into most versions of minimalism. The central idea in Kayne's analysis was that English prepositions resemble verbs in that their complements are always accusative and able to undergo movement. French prepositions, on the other hand, have inherently Case-marked complements (in the sense defined in §2.3.2), which cannot undergo movement independently of the preposition. So the parameter distinguishing French from English is connected to the differential lexical properties of prepositions in the two languages. The existence of inherent Case in French is morphologically signalled by the contrast between dative and accusative 3rd-person clitic pronouns, for example, *le* (acc) vs. *lui* (dat.); English has no comparable contrast. (Icelandic distinguishes accusative and dative complements, and has Preposition-stranding of some types, as mentioned above; see Kayne (1984: 117) for discussion).

PEI French behaves like English in allowing Preposition-stranding. (26) illustrates this in a *wh*-question, a relative clause and a passive:⁹

- (26) a. Où ce-qu'elle vient de __?
 where that she comes from
 'Where does she come from?'
 (King 2000: 136, (5))
- b. Ça, c'est le weekend que je me souviens de __
 That it is the weekend that I me remember of
 'That's the weekend that I remember.'
 (King 2000: 136, (6))
- c. Robert a été beaucoup parlé de __ au meeting.
 Robert has been much talked of at-the meeting
 'Robert was talked about a lot at the meeting.'
 (King 2000: 141, (32))

⁹ King points out (138) that Preposition-stranding has been observed in Montreal French (by Vinet (1984: 239)), but the phenomenon is much more restricted in that variety, according to King's account of Vinet's observations. Roberge (1998; 1999) surveys a range of Canadian varieties of French and observes that Alberta French is intermediate between Montreal French and PEI French.

Data like this indicate that PEI French has what appears to be English-style Preposition-stranding. Thus, PEI French, thanks to the very extensive contact with English, has developed a parametric option which is lacking in Standard French, and not fully instantiated in Montreal French; see note 9.

King insightfully and convincingly relates PEI French Preposition-stranding to the fact that PEI French has borrowed a number of English prepositions. Some of these are illustrated in (27):

- (27) a. Ils avont layé off du monde à la factorie.
 They have laid off some people at the factory
 ‘They have laid off people at the factory.’
 (King 2000: 142, (39))
- b. Il a parlé about le lien fixe.
 he has talked about the link fixed
 ‘He talked about the fixed link.’
 (King 2000: 143, (41))

These prepositions also allow stranding:

- (28) a. Qui ce-qu’a été layé off __?
 who that has been laid off
 ‘Who has been laid off?’
 (King 2000: 142, (40))
- b. Quoi ce-qu’il a parlé about __?
 what that he has talked about
 ‘What did he talk about?’
 (King 2000: 143, (42))

King argues that ‘the direct borrowing of English-origin prepositions has resulted in the extension of a property of English prepositions, the ability to be stranded, to the whole set of Prince Edward Island prepositions’ (147). If the option of stranding is genuinely a lexical property of prepositions, as roughly sketched in our remarks on Kayne (1984) above, then we might expect that option to be borrowed with the English prepositions, although PEI French does retain an accusative–dative distinction in pronouns, as King (2000: 64, Table 5.2) shows. So here we have a fairly clear case of direct contact: at some point in the history of PEI French, elements from an alien grammatical system – English prepositions – were borrowed and this affected the parameter governing Preposition-stranding. (Acquirers seem to have generalized the input based on the English prepositions; we will return to this notion of ‘generalization of the input’ in §3.5.) This is

borrowing, rather than in any sense imperfect learning of French by native speakers of French; nor does it reflect imperfect learning of English by French speakers.

In fact, PEI French appears to allow Preposition-stranding in contexts where English does not. Hornstein and Weinberg (1981) observed that examples like the following are unacceptable for most English speakers:¹⁰

(29) *Who did Pugsley give a book yesterday to __?

PEI French appears to allow examples equivalent to (29):

- (30) a. Quoi ce-que tu as parlé hier à Jean de __?
 what that that you have spoken yesterday to John of
 ‘What did you speak yesterday to John about?’
 (King 2000: 146, (53))
- b. Quoi ce-que tu as parlé hier de __ à Jean?
 what that that you have spoken yesterday of to John
 ‘What did you speak yesterday about to John?’
 (King 2000: 146, (57))

King relates this to the independent fact that ‘French does not have the strong adjacency requirements found in English’ (147).¹¹

The Prince Edward Island case is, as we said, a clear case of direct contact affecting the trigger experience. Reanalysis is relevant to the extent that the structure [_{PP} P DP] changes its properties as DP becomes extractable from PP. It is hard to evaluate the role of strong P-ambiguity here, as the pied-piping option is apparently only available with certain prepositions; with *de*, it is not found: **De quel enfant as-tu parlé?* (‘About which child did you speak?’) is not good, but in other cases it is possible: *Pour quelle raison qu’il a*

¹⁰ Hornstein and Weinberg suggest that this is because the verb must c-command the stranded preposition, but that in (29) the PP is ‘extraposed’ outside of VP and to its right, as its position relative to the adverb *yesterday* shows, and therefore the PP is not c-commanded by the verb. The relevant parts of the structure would thus be approximately as in (i):’

(i) [... [_{VP} V] yesterday] PP

The definition of c-command was given in §1.4, (90).

¹¹ It could in fact be connected to the possibility of raising a participle to a slightly higher structural position than that occupied by English participles, allowing the verb a wider range of c-command possibilities, roughly in line with what is suggested in note 10. On raising of French participles, see Pollock (1989: 417). We mentioned movement of non-finite verbs in French in §1.3.1, note 10.

parti?/Quelle raison qu'il a parti pour? ('For what reason did he leave?/What reason did he leave for?') (Ruth King, p.c.). It is clear, though, that strings with Preposition-stranding crucially affected the PLD at the time of contact, leading PEI French to change the value of the Preposition-stranding parameter, thus creating a difference with Standard French.

3.3.2. Cue-driven parameter-resetting

Let us now turn to the second way of making the trigger experience ambiguous and opaque: the cue-based approach. Lightfoot (1999) and Dresher (1999) argue that learners use input forms, i.e. pieces of PLD, as cues for setting parameters. The trigger in this case is not sets of sentences but fragments of utterances (partial structures) (cf. also Fodor (1998) on the potential importance of fragments of sentences for parameter-setting). For Dresher (1999) each parameter has a marked and a default setting, and comes with its cue, as part of the UG specification of parameters. For example, Dresher (30ff.) proposes that there is a parameter determining whether a given language's stress system is quantity-sensitive (QS) or not (Q(uality)I(n)sensitive)). English, for example, is QS, in that the basic stress rule states that the penultimate syllable is stressed if heavy; otherwise the antepenult is stressed (cf. *Cánada*, with a non-heavy CV penult, as opposed to *Vancóu:ver*, with a heavy CV: penult). Thus the heaviness (or quantity) of a syllable plays a role in determining stress-assignment. Not all languages have quantity-sensitive stress-assignment; QS thus represents a value of a particular parameter. The parameter in question is formulated as follows:

(31) *Quantity (in)sensitivity*

- a. *Parameter*: The language {does not/does} distinguish between light and heavy syllables ...
- b. *Default*: Assume all syllables have the same status (QI).
- c. *Cue*: Words of *n* syllables, conflicting stress contours (QS).
(Dresher's (7): 31)

Dresher (1999: 31) points out that:

In QI systems all words with *n* syllables should have the same stress contour, since they are all effectively equivalent. Taking quantity insensitivity to be the default case, a learner will continue to assume that stress is QI until it encounters evidence that words of equal length can have different stress contours.

Dresher goes on to point out that there is abundant evidence that English is not QI, and so the learner quickly sets this parameter to the QS value. Thus we see how the unmarked value requires no evidence, and the marked value is associated with the cue. We will look more closely at the question of the marked and default settings of syntactic parameters in the next section.

Lightfoot (1999: 149), however, takes a much stronger view and argues that ‘there are no independent “parameters”’; rather, some cues are found in all grammars, and some are found only in certain grammars, the latter constituting the points of variation’. He illustrates this approach with the loss of V-to-T movement in ENE. He assumes that the NE situation whereby tense and agreement morphology is realized on V is the default (Lightfoot says that this is a morphological rule, but we can continue to think of it as an instance of Agree; see §1.4.1), and so the V-to-T grammar needs the cue [_T V] (‘_I V’ in his notation). Lightfoot suggests that the main expression of the cue (where his notion of ‘expression’ is like our notion of P-expression in being a structure which requires the cue in order to be grammatical) to be subject-verb inversion, as in (75) of Chapter 1, repeated here as (32):

(32) What **menythe this pryste?**

What does this priest mean?

(1466–7: Anon., from J. Gairdner (ed.), 1876, *The Historical Collections of a London Citizen*; Gray 1985: 11; Roberts 1993a: 247)

This cue was perturbed by three factors: (i) the reanalysis of modals as T-elements (see §2.1); (ii) the development of dummy *do* in the sixteenth century, also a T-element (we briefly mentioned this in our discussion of the loss of V-to-T in §2.1; it may well have been the same change as that affecting modals); (iii) the loss of V2, which clearly took away many environments in which verb-subject order had formerly been found. He concludes:

with the reanalysis of the modal auxiliaries, the increasing frequency of periphrastic *do*, and the loss of the verb-second system, the expression of _I[V] in English became less and less robust in the PLD. That is, there was no longer anything very robust in children’s experience which had to be analysed as _I[V], which *required* V to I, given that the morphological I-lowering operation was always available as the default. (Lightfoot 1999: 164)

This account is very similar to the one we proposed in §2.1, with the notion of cue playing the role of our notion of P-expression. Lightfoot also points

out that ‘weak’ verbal agreement inflection is a precondition for the change (‘the *possibility* of V to I not being triggered first arose in the history of English with the loss of rich verbal inflection’ (164)), although he does not explicitly say that this inflection is the expression of a cue, which would be the equivalent in his terminology of our claim in §2.1 that the relevant agreement morphology expresses the parameter.

The similarities between Lightfoot’s cue-based account of the loss of V-to-T and the one we gave in §2.1, are, as we see, very close. In fact, the definition of ‘trigger’ in (20b), in making reference to ‘a substring of the input text S’, is equivalent to the Lightfoot/Dresher notion of cue. However, in this sense, cues cannot be identified with parameters: parameters are abstract properties of grammars, features of part of an individual’s mental representation. Although the notion of cue is useful, it must be kept distinct from the notion of parameter.

A further point is that Lightfoot’s cue-based approach is too unconstrained: if there is no independent definition of cues, then we have no way of specifying the class of possible parameters, and hence the range along which languages may differ, synchronically or diachronically. It is, however, possible to maintain that parameters can be independently defined and that learners also make use of cues provided by the input (this is closer to Dresher’s view); if we do this we do not run into this difficulty. So, it seems reasonable to take the view that cues, i.e. triggers as defined in (20b), are provided by the input; parameters are specified by UG and are set by the learner on the basis of the interaction of cues/triggers and UG (and internal properties of the learner – see §3.4 and §3.5). Construed this way, the Lightfoot/Dresher view is essentially the one I have been presenting here, as the close similarities in the analysis of the loss of V-to-T show.

However, there is a difference: Lightfoot has no account for the shift in the cue. He says:

this model ... has nothing to say about why the distribution of cues should change. This may be explained by claims about language contact or socially defined speech fashions, but it is not a function of theories of grammar, acquisition or change – except under one set of circumstances, where the new distribution of cues results from an earlier grammatical shift; in that circumstance, one has a ‘chain’ of grammatical changes. One example would be the recategorization of the modal auxiliaries ..., which resulted in the loss of V to I.

(Lightfoot 1999: 166)

A morphologically-based approach like that sketched earlier can, on the other hand, explain the change in the cue/P-expression in terms of morphological loss. Let us now return to that account.

3.3.3. *Morphologically-driven parameter-resetting*

It is worth making two points here. First, we observed in §2.1 that there was a strong P-ambiguity in the analysis of very simple positive declarative sentences like *John walks* in sixteenth-century English, as schematized in (33) (repeated from (13) of Chapter 2):

- (33) a. John [_T walks] [_{VP} ... (walks) ...]
 b. John T [_{VP} walks]

We saw that Lightfoot also makes this observation. (33a) represents the conservative structure with V-to-T movement, and (33b) the innovative structure without V-to-T movement. The proposal was that the conservative system was preferred as long as there was a morphological expression of V-to-T movement through the agreement system. This was stated in terms of the following postulate, linking agreement marking on the verb to V-to-T movement (repeated from (18) of Chapter 2):

- (34) If (finite) V is marked with person agreement in all simple tenses, this expresses a positive value for the V-to-T parameter.

We then proposed that the loss of much verbal agreement, particularly plural endings in both simple tenses, led to the loss of the morphological expression of the V-to-T parameter and thus to a reanalysis of (33a) as (33b) with the concomitant change in the value of the V-to-T parameter. Since (33a) contains an occurrence of movement, V-to-T movement, missing from (33b), it must have at least one more formal feature than (33b). Hence, by (22), it is more complex than (33b). So here the crucial factor creating ambiguity and opacity in the PLD is the erosion and loss of certain endings, something I take to be a morphological property. (In fact, it is more than likely that it is ultimately phonological; see Lass (1992: 134ff.)) This solves the Regress Problem and gives an account of what changed in the P-expression/cue-expression, unlike Lightfoot's analysis.

The second point concerns weak P-ambiguity, as defined in (21c). As a comparison of the definition of strong P-ambiguity in (21b) and that of weak P-ambiguity in (21c) shows, the essential difference lies in the fact that

a weakly P-ambiguous string triggers neither value of a parameter. This notion can be relevant to understanding certain aspects of change in that an independent change can render a former trigger weakly P-ambiguous, i.e. render it irrelevant for triggering some value of a parameter that it triggered prior to the independent change. The loss of V-to-T movement in ENE exemplifies this. The reanalysis of the modals and *do* as functional elements merged in T had the consequence that examples like (35) (repeated from (14) of Chapter 2) no longer triggered anything, i.e. they were weakly P-ambiguous:

- (35) a. I may not speak.
b. I do not speak.

Prior to reanalysis of modals and *do* as T-elements, such examples provided an unambiguous trigger for a positive setting for the V-to-T parameter, in that modals and *do* were verbs (with plural agreement marking) which moved to T and expressed the morphological trigger for V-to-T movement. Once modals and *do* are merged in T, such sentences become weakly P-ambiguous in relation to the V-to-T parameter in that they are compatible with either value of the parameter. As such, an important, and frequently occurring, kind of trigger for the positive setting of the V-to-T parameter is lost. Weak P-ambiguity may be relevant to understanding certain changes in this way. (We will discuss weak P-ambiguity more in §5.3 and §5.4.)

Another change discussed in Chapter 2 illustrates further how morphological change may affect the PLD in such a way as to create ambiguity and opacity in triggers and hence abductive change. This concerns the loss of so-called recipient passives in thirteenth-century English, as discussed in §2.3.2. There we schematized the crucial reanalysis as in (36) ((43) of Chapter 2):

- (36) [_{CP} Him+DAT [_{TP}[_T was] [_{VP} v [_{VP} helped (him+DAT)]]]]>
[_{TP} He+NOM [_T_{uφ}] was] [_{VP} v [_{VP} helped (he+NOM)]]]

We treated this reanalysis as directly caused by the loss of dative-case morphology, i.e. by the loss of any morphological distinction between morphological accusative and morphological dative case. This led, we proposed, to a parametric change in the nature of *v*, in that it henceforth had a new uninterpretable Case feature which was available in all transitive clauses, with the consequence that all inherently Case marked non-subject arguments were lost.

We tied the reanalysis in (36), and the associated parametric change in the Case-features of *v*, to the loss of case-marking distinctions among complements. This is directly supported by Allen's (1995) detailed account of the breakdown of case morphology in English, as we saw. Let us now consider the effect of the loss of dative-case morphology on the relevant PLD in more detail. Consider a variant of (36) with a non-pronominal argument, which therefore after the loss of dative-case morphology is fully ambiguous as to which Case-feature it bears. We also make the argument singular, so that verbal agreement cannot tell us whether it is Nominative. The string in question is thus *the man was helped*. One could assume that the situation is straightforward here: if there is no dative morphology there is no abstract Dative Case. However, two points militate against this simple assumption. First, Allen (1995: 351ff.) shows that direct passives (i.e. those with Nominative subjects) appear in dialects where the dative/accusative distinction still remains. Second, the relation between abstract Case and concrete case is not usually one-to-one; instead, it normally has the form of a one-way implication, viz. (see Kayne (1984: 116–17) on this):

- (37) If a DP has morphological dative case, then the grammar has abstract Dative Case.

This is in fact the simplest statement of the relation between morphological case and abstract Case that one can postulate, assuming the existence of any kind of abstract Case at all. What this implies is that as long as there was morphological dative case there could be no ambiguity at all regarding these constructions. However, it says nothing about the situation once the case morphology has been lost. One might conclude that the string in question is weakly P-ambiguous, since it provides no unambiguous information regarding the parametric property of *v*. But the very fact that, thanks to the one-way implication in (37), there could be an abstract DAT, shows that the structure must be strongly P-ambiguous (since the presence of DAT implies one feature make-up for *v*, while the absence of DAT implies that active *v** has uninterpretable φ -features; see §2.3.2).

The structural ambiguity is partially represented in (38), bearing in mind that English was a V2 language at this time, and supposing that our example is a main clause:¹²

¹² See note 17 of Chapter 2 on the reason for assuming that these clauses are V2-clauses, i.e. CPs.

- (38) [_{CP} The man+DAT/NOM [_C was] [_{TP} [_T_[u φ] (was)] [_{VP} v [_{VP} helped
(the man+DAT/NOM)]]]]

Example (38) shows that the clause is a CP, and that *the man*, which we are taking to be ambiguously Nominative or Dative, occupies SpecCP. What is left unclear in (38) is the nature of SpecTP and the way in which T's uninterpretable φ -features are eliminated where *the man* is DAT. It is usually assumed that SpecTP must be filled, and whatever fills this position must be able to Agree with some feature in T, i.e. that T has an EPP feature (see §2.5). Where *the man* is NOM, we can assume that it moves through the SpecTP position on its way to SpecCP, thereby satisfying the EPP. And of course, *the man's* NOM-feature Agrees with T's uninterpretable φ -features and so these uninterpretable features are able to be eliminated from the representation and the NOM-feature is valued. On the other hand, where *the man* is DAT, it has no feature which can Agree with T. (This is clear where we have a plural argument, as there is no agreement in number between a dative argument and the verb, i.e. between a DAT DP and T's φ -features; see Allen (1995: 70ff., 142ff.) for discussion.) So, if *the man* is DAT, it is unable to move through SpecTP on its way to SpecCP as it cannot Agree with any feature of T.¹³ Therefore the EPP must be satisfied in some other way in this situation. There are two options as to what can fill SpecTP. On the one hand, we can postulate an expletive null subject (i.e. *pro*), an element which can also bear NOM and thus Agree with T (see §1.2.1 on expletive, i.e. non-referential, *pro*). The other option is to assume 'massive movement', in the sense introduced in §2.5, of either vP or VP into SpecTP (see examples (83–8) in Chapter 2). In order for massive movement of this kind to satisfy the requirement that the element in SpecTP Agree with some feature of T, we have to assume that the moved category may contain the element which Agrees with T. In the present case, we may assume, following Baker, Johnson, and Roberts (1989), that the passive marker itself is a nominal element capable of bearing a Case feature which can Agree with T. (This idea is updated in the context of massive

¹³ Chomsky (2000: 128) suggests that the very similar dative subjects in Icelandic might Agree for a Person feature with T, but, as mentioned in note 17 of Chapter 2, there is no evidence that proposed datives were subjects in passives in OE. It thus seems correct to take the DAT DP to occupy SpecCP and to have not moved through SpecTP. So, at least in OE and ME, there is no Agree for person features in this construction.

movement by Richards and Biberauer (2005.)) So in fact our example in (38) really manifests a strong P-ambiguity between (39a) and either (39b) or (39b') (this is a more elaborate version of the reanalysis given in (36) above and in (43) of Chapter 2):

- (39) a. [_{CP} The man+NOM [_C was] [_{TP} (the man+NOM) [_{T[u_φ]} (was)] [_{vP} v [_{vP} helped (the man+NOM)]]]]
 b. [_{CP} The man+DAT [_C was] [_{TP} *pro*+NOM [_{T[u_φ]} (was)] [_{vP} v [_{vP} helped (the man+DAT)]]]]
 b'. [_{CP} The man+DAT [_C was] [_{TP} [_{vP} v [_{vP} helped+NOM (the man+DAT)]] [_{T[u_φ]} (was)] (vP)]]

Example (39a) is quite unproblematic: *the man* Agrees for Nominative with T, and moves through SpecTP on its way to SpecCP, satisfying the requirement for an element in SpecTP, which Agrees with T. In (39b), *the man* bears the interpretable DAT feature which does not need to Agree with anything. It moves in one step to SpecCP. SpecTP is filled by expletive *pro*, which is NOM and so Agrees with T. In (39b'), *the man* behaves exactly as in (39b). However, vP raises to SpecTP and the passive marker on *helped*, which is contained in vP, Agrees with T.

The ambiguity in (39) is created by the loss of dative morphology. Given (37), where there is dative morphology, one of the options (39b) or (39b') is the only one. Once dative case is lost, (39a) becomes available. (39a) involves two movements of *the man+NOM*: first to SpecTP and then to SpecCP. (39b) involves just one movement of *the man+DAT* to SpecCP, and insertion of expletive *pro* in SpecTP. (39b) therefore appears to be less complex than (39a). We could attempt to introduce some further complexity cost associated with the postulation of expletive *pro*, but probably the best course of action is to assume that expletive *pro* is not relevant here and that instead the correct representation of the Dative option involves a structure with massive movement like (39b'). Since this structure involves copying of all the vP-internal material, along with all its formal features, this structure will be more complex than (39a). In that case, the definition of complexity in (22) gives the right results. As a consequence of this reanalysis and the associated parametric change, v's feature make-up was changed as described in §2.3.2 with the consequences outlined there. Without the loss of dative-case morphology, Inertia ensures that the structure remained as (39b') (not (39b) if we rule out expletive *pro*).

A final illustration of the role of morphological change causing the PLD to change comes from our discussion of complementation in Latin and Romance in §2.4. Among various other changes discussed there, we suggested that Latin infinitives were associated with a T-position which was able to Agree for Accusative with an argument in the subordinate clause. This gives rise to the accusative + infinitive construction in Latin (or, more precisely, one variant of it – see the discussion in §2.4 and below), as in (repeated from (56a) of Chapter 2):

- (40) Dico te venisse.
 I-say you-Acc come-perf-infin
 ‘I say that you have come.’

It was tentatively proposed that Latin non-finite T lost this capacity when the tense/aspect forms of the infinitive such as perfect *venisse* were lost. In line with (34) and (37) above, let us state this as a one-way implicational statement, as follows:

- (41) If T[-finite] has an Accusative feature, then it shows inflectional distinctions marking tense/aspect.

This means that after loss of forms such as *venisse*, the unmarked (formerly present) form of the infinitive was no longer associated with an Accusative-bearing T. Nevertheless, examples like (42), with an Accusative subject of the complement clause and the unmarked form of the infinitive, would have been possible:

- (42) Dico te venire.
 I-say you-Acc to-come

We noted in §2.4 that this construction was ambiguous in Classical Latin in that the Accusative feature of the subject of the infinitive *te* could originate either in the infinitival T, or in the main-clause v^* . (The evidence for this comes from the two attested passive constructions in (53a) and (54) of Chapter 2.) After the loss of the tense/aspect forms of the infinitives, examples like (42) became less ambiguous than previously, in that they became unambiguously English-style accusative + infinitives, with the Accusative subject agreeing with v^* in the superordinate clause. This possibility was, however, then ruled out by a change in the value of Parameter G, itself connected to the development of Romance-style prepositional complementizers, and so Accusative subjects of infinitives were eliminated in general. So the loss of morphology played a role in eliminating

the accusative + infinitive construction in the development from Latin to Romance, although in this instance by eliminating an ambiguity rather than creating one.

3.3.4. *Conclusion*

Here we have tried to apply the solution to the logical problem of language change proposed in the previous section to actual cases of change discussed in the earlier chapters. If contact, cues, or morphology cause changes in PLD, then P-ambiguity and consequent opacity of one representation result, leading to abductive reanalysis and associated parametric change. We also saw that there is often an implicational relation between a morphological trigger and a parameter value. We will return to this last point briefly in the next section.

The main question that has been begged throughout the discussion is that of the definition of opacity. This, along with the closely-related notion of markedness, is the subject of the next section.

3.4. **Markedness and complexity**

The purpose of this section is to connect complexity as defined in (22) above to markedness, and thereby arrive at a basis for defining the marked and unmarked values of parameters with a view to formulating parameters along the lines of Dresher's proposal illustrated in (31) above. The concept of complexity is closely related to that of markedness. Here I will discuss an approach to determining in general the marked and unmarked values of parameters which correlates marked parameter values to the relative opacity or complexity of representations or derivations. The idea is that marked settings are associated with opaque, that is relatively complex, constructions.

3.4.1. *The concept of markedness*

The concept of markedness originated in Prague School phonology, apparently with Trubetzkoy, and was taken up by Jakobson (1941). (The history

of the concept is described in detail in Battistella (1996: 19ff.)) The basic idea can be stated as follows: given a binary opposition, the two terms of the opposition may stand in a symmetric relation or in an asymmetric one. In the former case, we say that the terms are equipollent; in the latter, we refer to one term as the marked one and the other as the unmarked one. The asymmetry lies in how the absence of specification of the terms is interpreted: the absence of the marked term implies the unmarked term, but the *absence of the unmarked term does not on its own imply the marked term*. In other words, all other things being equal we assume the presence of the unmarked term; the presence of the marked term requires something special, however, i.e. some kind of ‘mark’. This asymmetric formulation can be maintained independently of the nature of the terms involved, the nature of the asymmetry, or the correlates of the asymmetry in some other domain. It is often supposed that the marked term is associated with relatively greater complexity; this is arguably inherent in the idea of it requiring an extra ‘mark’. For example, Cinque (1999) proposes a series of markedness conventions for the features associated with various functional heads in his analysis of clause structure. He states that marked features are ‘more restricted [in] application, less frequent, conceptually more complex, expressed by overt morphology’ (128), while unmarked features are in each case the opposite.

We can illustrate the essential asymmetry that characterizes markedness relations with phonological distinctive features. A phonological opposition, for example that of voicing, can be thought of as an equipollent opposition between [+Voice] and [–Voice] or an asymmetric opposition between [mVoice] and [uVoice]. (Here and below, ‘u’ before the name of a feature means ‘unmarked’, not ‘uninterpretable’ as in the specifications of formal syntactic features in earlier sections; ‘m’ means ‘marked’.) Chomsky and Halle (1968, Chapter 9) discuss markedness in relation to the phonological distinctive-feature system they proposed; the approach to markedness and complexity adopted here is largely inspired by their discussion. Where the opposition is equipollent, if a segment is not [+Voice] then it is [–Voice] and vice-versa. But where the opposition involves a markedness asymmetry, markedness conventions and perhaps other statements are required to determine the value of the coefficient of a feature (Chomsky and Halle 1968: 403ff.). Moreover, there need not be a straightforward relation between the *u/m* values and the *+/-* values; for example, Chomsky and Halle (1968: 406) proposed that [uVoice] is [–Voice] if the segment is

[−sonorant], but [+Voice] if it is [+sonorant]. Marking conventions of the type first put forward by Chomsky and Halle imply that the underspecified feature [Voice] can ‘default’ to a given value under various circumstances, something impossible in the case of the equipollent +/− opposition.¹⁴ In turn, this leads to the possibility of an ‘elsewhere convention’, a notion going back to the Sanskrit grammarians of Indian antiquity (see Kiparsky (1973)): in the absence of specification, the unmarked feature value is assumed, while a marked value requires positive specification, and therefore a longer description. For Chomsky and Halle, the markedness asymmetry relates to the evaluation metric they propose for determining the relative simplicity of rule systems: ‘the unmarked value of a feature was cost-free with respect to the evaluation metric, while the marked values were counted by the metric’ (Battistella 1996: 75). The correlates of the asymmetry were stated by the marking conventions (Chomsky and Halle (1968: 404–7) propose thirty-nine of these), which are intended to capture aspects of the intrinsic content of distinctive features. The correlates of markedness in the distinctive-feature system include: cross-linguistic frequency of unmarked terms (all languages have voiceless obstruents, but not all have voiced ones: note how in implicational universals, the marked value of an opposition entails the unmarked one (see Croft (2003) for discussion)); unmarked terms appear earlier than marked ones in language acquisition and are lost later in language deficits (this was first proposed by Jakobson (1941)); and the fact that unmarked values emerge under neutralization in certain positions, for example, the coda of a syllable or the end of a word (for example, final-obstruent devoicing is cross-linguistically very common, while obstruent voicing in this context is relatively rare). Kenstowicz (1991: 61–4) discusses these points in more detail in relation to phonology.

3.4.2. *Markedness and parameters*

Since we take parameters to have binary values (see §1.1 for general discussion, and note that all the examples of parameters we have discussed have been formulated in a strictly binary fashion), we can in principle apply

¹⁴ The idea that there is no single unmarked value for a feature, but that this may depend on other features, represents an important difference between Chomsky and Halle’s approach and the Prague School approach to markedness.

markedness logic to the opposition between these values; in other words, we can treat the binary opposition between the two values of a parameter as an asymmetric one in the sense described above. This has in fact been suggested in various places ever since the earliest formulations of principles-and-parameters theory (see Chomsky (1981), and the discussion in Battistella (1996: 82ff.)).¹⁵ If we do this, we have to answer three questions: (i) what is the nature of the features involved in the asymmetric relation? (ii) what is the nature of the asymmetry? (iii) what are the correlates of the asymmetry in other domains?

Regarding question (i), it suffices for now to simply treat the features in question as the values of a parameter; giving a fuller answer requires a proper statement of the form of parameters, something we have yet to do. In the next section, I will attempt a general characterization of parameters which will provide a more substantive answer to this question, and thereby facilitate the statement of parametric marking conventions similar to those introduced into phonological theory in Chomsky and Halle (1968, Chapter 9).

We could answer question (ii) in terms of the definition of complexity given in (22), which we repeat for convenience:

- (22) Given two structural representations R and R' for a substring of input text S, R is simpler than R' if R contains fewer formal features than R'.

The nature of the asymmetry between the parameter values lies in the complexity of the structures generated by the grammars determined by the different values. The unmarked value of a parameter determines a grammar which generates simpler structures than those generated by the marked value. In the next section, we will suggest that parameter interactions give rise to a slightly more complex and interesting situation than this.

¹⁵ Chomsky's discussions of markedness here and in *Knowledge of Language* (1986), make use of the distinction between the 'core grammar' and the 'periphery' in various ways. I am not assuming this distinction, as seems to be more in line with Chomsky's assumptions in his more recent work on the Minimalist Program, where this distinction should no longer play a role. (This is what I take to be the implication of Chomsky's remark that it 'should be regarded as an expository device, reflecting a level of understanding that should be superseded as clarification of the nature of linguistic inquiry advances' (Chomsky 1995: 163, note 3)). As the text discussion will make clear, I follow Chomsky's thinking in taking markedness to impose a preference structure on the parameters of (core) grammar for the language acquirer.

Question (iii) brings us back to the issue of most concern here. At least some of the correlates of the markedness asymmetry between parameter values lie in syntactic change: since abductive reanalysis and parametric change arise through P-ambiguity and opacity/complexity of the trigger, with the less complex structure being preferred, then – all other things being equal – parametric change will be in the direction of unmarked values.¹⁶ We expect to find correlates in language acquisition (for example, marked values being harder to acquire and hence acquired later) and typology (marked values being less cross-linguistically frequent). I will not comment further on language acquisition, given the conclusion of §3.1 that it is hard to observe the postulated connection between acquisition and change. I will come back to the relationship between markedness and typology in the next section.

The approach to complexity in (22) comes remarkably close to the notion of ‘value’ put forward by Chomsky and Halle (1968: 334): ‘The “value” of a sequence of rules is the reciprocal of the number of symbols in its minimal representation’. Taking the relevant symbols to be formal features, which are the important symbols in the syntactic representation in the theory of syntax being assumed here, the definition in (22) would make it possible to value syntactic derivations just as Chomsky and Halle propose valuing phonological derivations. We did essentially this in our discussion of the role of complexity/opacity in abductive reanalysis in the previous section.

A further point arises from this. Roberts and Roussou (2003: 210–13) derive a series of markedness hierarchies from the definition of complexity in (22). Here I give a simplified version of their hierarchy:

(43) Move > Agree > neither

Here ‘>’ means ‘is more marked than’. So a category set to a parameter value which requires movement is more marked than one which merely causes an Agree relation, which is in turn more complex than one which has neither property. This follows straightforwardly from the feature-counting idea: in order to give rise to movement, a category must have both

¹⁶ Of course, we do not want only this kind of change to be possible: change from unmarked to marked must be allowed somehow. This point will be dealt with in the next section.

uninterpretable formal features and the movement-triggering (EPP) feature. In order to give rise to an Agree relation, a category need only have an uninterpretable formal feature. Finally, in order to trigger neither operation, a category should lack both EPP and uninterpretable formal features.¹⁷ This approach is developed in some detail in Roberts and Roussou (2003, Chapter 5), and we will consider some of its implications below.

3.4.3. *The Subset Principle*

Other approaches to determining the marked and unmarked values of parameters have been put forward. One important and influential proposal was the Subset Principle of Berwick (1985). This states that ‘the learner selects the grammar that generates the smallest possible language that is compatible with the data’ (Manzini and Wexler 1987: 425). The interest of the Subset Principle is that it relates to an important and fairly well-established aspect of language acquisition: the fact that language acquirers do not have access to negative evidence. In other words, language acquirers are not presented with ungrammatical sentences which are marked as such. As Guasti (2002: 4) puts it, ‘negative evidence is not provided to all children on all occasions, is generally noisy, and is not sufficient . . . Children have the best chance to succeed in acquiring language by relying on positive evidence [emphasis omitted – IGR], the utterances they hear around them – a

¹⁷ One might wonder whether a category with three uninterpretable features is more marked than one with one uninterpretable feature and an EPP feature. This is predicted by (22), but is not consistent with the hierarchy in (43). (43) should be understood as holding in relation to a given feature: in that case, Move-F will always be more complex, and therefore more marked, than Agree-F, for any F, as Move requires the EPP feature in addition to F. Chomsky (2005c) introduces the possibility of Move occurring independently of Agree, possibly triggered by a further kind of feature known as an Edge Feature (EF). EF-triggered movement is characteristic of wh-movement, topicalization and focalization, movements which typically target the ‘left periphery’ of the clause. It is possible, and would follow from (22), that this type of movement is less marked than that triggered by EPP where Agree is involved. See the discussion of how grammars may innovate marked properties in §3.5 below.

resource that is abundantly available'. (We touched on the absence of negative evidence in our discussion of the poverty of the stimulus in Chapter 1).

Because they only have access to positive evidence, acquirers are, other things being equal, at risk of falling into a 'superset trap.' This can happen if acquirers posit a grammar which generates a language which is a superset of the language generated by the actual grammar in their environment, in the sense that it contains no examples that are incompatible with the PLD to which the children are exposed, but it generates examples that are incompatible with the target grammar. (This situation would correspond to the schema for abductive change in (13) of §3.2 above, with G_1 a subset of G_2 .) If children only have access to positive evidence, they will never hear any example which causes them to 'retreat' from the superset grammar. Thus they may posit a grammar which is incompatible with the target, and recall that it is a standard assumption in work on language acquisition that this does not happen; this is what underlies the empirical force of the Inertia Principle – see the discussion in §3.2 above.

In order to rule out the risk of superset traps, the Subset Principle is proposed as a condition on language acquisition. The Subset Principle, as just given in the quotation from Manzini and Wexler (1987), forces children to hypothesize the grammar which generates the smallest language compatible with the trigger experience. In this way, it is argued, they do not run the risk of falling into superset traps.

The notion of markedness which derives from this then is that marked parameter values will generate bigger languages. The null-subject parameter may serve as a (slightly artificial) example. As we saw in §1.1.1, null-subject languages allow a definite, referential pronoun subject to be dropped in finite clauses, while non-null-subject languages do not:

- (44) a. Parla italiano.
b. *Speaks Italian.

On the other hand, null-subject languages typically allow the pronominal subject to be expressed, just like non-null-subject languages:

- (45) a. Lui parla italiano.
b. He speaks Italian.

(As we saw in connection with examples (14) and (15) in Chapter 1, there are interpretative differences between null-subject and non-null-subject

languages where subject pronouns are expressed in the former; I will gloss over these for the purposes of illustration of the Subset Principle, however). Thus the grammar of Italian generates a larger set of strings than that of English. In other words, non-null-subject languages are a subset of null-subject languages. This can be illustrated as follows:

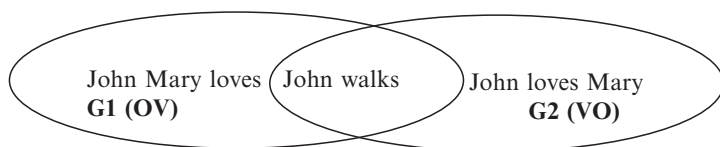
(46)



The Subset Principle therefore implies that the positive setting of the null-subject parameter is more marked than the negative setting.

One empirical problem that one could raise here is the evidence of early null subjects in L1 acquisition of non-null-subject languages. However, as we saw in §3.1, this phenomenon probably is not related to a ‘missetting’ of the null-subject parameter at an early stage of language acquisition, and so the objection does not hold. A much more serious problem with the above line of reasoning emerges if we consider the various parameters we have put forward in our discussion: verb-movement parameters (both V-to-T and V2), the negative-concord parameter, the wh-parameter, and word-order parameters all define intersecting grammars. That is, in each case, one setting of the parameter allows one type of structure S and disallows another type S', while the other setting allows S' and disallows S. This is clearest in the case of word-order parameters: one setting of parameter F1, for example, allows VO and at the same time disallows OV, while the other setting has just the opposite effect. The intersection relation can be illustrated as follows:

(47)



The material in the intersection is weakly P-ambiguous in relation to the parameter P in question in terms of the definition in (21c), while the material in the complement of the intersection expresses the value of P.) The same exercise could be repeated for the verb-movement parameters, the negative-concord parameter, and the wh-movement parameter. In fact, it can also be repeated for the null-subject parameter, to the extent that it is true that null-subject languages do not have overt expletives:

(48)



As Battistella (1996: 113) points out: '[i]f markedness relations obtain between parameters that are not in a subset relation, they must be accounted for in some other way'. The above considerations seem to indicate that the Subset Principle is not a useful way of predicting markedness relations in general among parameters, since most – if not all – parameters define intersection relations of the kind seen in (47) and (48).

A further issue arises if we look at the Subset Principle in the diachronic domain. If non-null-subject languages are subsets of null-subject languages, then we expect a diachronic preference for change from the positive to the negative value of this parameter. We know that the null-subject parameter must have changed from positive to negative at some point in the history of Germanic (see the remarks on this in §3.1 above), and this change has certainly happened in the history of French and certain Northern Italian dialects, as we saw in §1.1.2. So this much is consistent with what the Subset Principle predicts. However, we also saw there that this parameter may have changed its value in the opposite direction in exactly these Romance varieties, with some question as to what may be the best analysis of contemporary French.

So we conclude, rather reluctantly, that the Subset Principle is not useful in providing the basis for determining the markedness of parameter values in cases like the above. The reluctance is due to the fact that the Subset Principle has the great conceptual merit of being firmly grounded in an important fact about acquisition: that children do not have access to negative evidence.

One area where the Subset Principle may be useful is in distinguishing between a grammar which allows genuine formal optionality and one which does not. Abstractly, a case of this type would be where G_1 allows an alternation between two constructions C_1 and C_2 while G_2 , thanks to a

different parameter-setting, does not. An example might be the difference between English and (Standard) French regarding Preposition-stranding or pied-piping. (This was discussed in Chapter 1, Box 1.5 and in the previous section.) English has the option of Preposition-stranding or pied-piping, while French only allows the latter. The French situation is illustrated by (25) above, which we repeat here:

- (49) a. Who did you speak to __?
b. To whom did you speak __?
- (25) a. *Qui as-tu parlé à __?
b. A qui as-tu parlé __?

The fact that English allows both options, while French only allows one of them means that the French parameter-settings generate a language which is a subset of the English one. We could, therefore, regard English as marked in relation to French in this respect. Of course, (49b) is characteristic of a relatively ‘high’ register as compared to the more colloquial (49a), but for the purpose of this illustration of the logic of the Subset Principle I abstract away from this; we will come back to the concept of the ‘social value’ of variants in §4.2.

The Subset Principle might also lie behind the phenomenon of ‘restriction of function’, whereby in one system a given operation applies more freely than in another. The more restricted grammar then produces a subset of the grammatical strings of the more liberal one. An example of this might be the restriction on OV orders to negative and quantified objects in fifteenth-century English which we mentioned briefly in §2.5.3. Given that OV order was an option with non-negative, non-quantified DPs in the earlier stage (i.e. ME from roughly 1200 to 1400), we have a situation where the fifteenth-century grammar only allowed OV for a particular class of objects and required VO elsewhere, while the earlier grammar allowed OV with any kind of object. Thus object-movement, assuming that is the correct analysis of this construction (see §2.5.4), was restricted in function. Here there may be a tension between concepts of markedness based on the Subset Principle and those based on feature-counting, since the more restricted variant requires more features.

We see then that the Subset Principle has a major conceptual advantage, being based on what appears to be an important fact about language acquisition: namely, that language acquirers do not make use of negative evidence. Its actual application to parametric systems may be somewhat restricted, since so many parameters appear to define languages in

intersection, rather than inclusion, relations. However, it may play a role in relation to true formal optionality, in predicting that such systems would be marked, and it may play a role in accounting for the diachronic phenomenon of ‘restriction of function.’

3.4.4. *Markedness and core grammar*

Another proposal, which was not intended to form the basis of a general account of the markedness of parameter values, was made by Hyams (1986: 156ff.). She took the view that markedness was a feature only of the ‘periphery’ of the grammar (in the sense briefly discussed in note 15 above). The null-subject parameter, however, is a property of core grammar, and so the question of the markedness of its settings does not arise. Nevertheless, on the basis of her observation of early null subjects in the production of children acquiring non-null-subject languages, she argues that the null-subject value is the more accessible value (Hyams 1986: 162–3) since null subjects do not require the costly process of lexicalization of pronouns (163). Hence children acquiring English begin with the assumption that it is a null-subject language, and the parameter is reset during the course of language acquisition to the negative value (on the basis of the evidence from modals and overt expletive subjects, as mentioned in §4.1). One could imagine that this would favour a tendency in language change in the direction of null subjects, but the general view now held amongst researchers on L1 acquisition is that early null subjects do not reflect a ‘missetting’ of the null-subject parameter, but rather some property of immature competence. For this reason we leave this proposal aside.

3.4.5. *Markedness and inflectional morphology*

At this point it is justifiable to ask what the advantages of an analysis of parameter values into marked and unmarked values might be. Aside from connecting syntactic change to the form of parameters, as we have done, one independent point has to do with the nature of language acquisition. Lasnik (1983) observes that there is an intrinsic connection between markedness in L1 acquisition and the question of indirect negative evidence. The notion of indirect negative evidence is discussed by Chomsky (1981: 8–9): although, as

we have suggested and as is the standard view amongst L1-acquisition researchers, children do not have direct negative evidence in the sense of not having access to the information that a given structure or string is ungrammatical, Chomsky suggests that indirect negative evidence may nevertheless be available in the case where some feature is expected by acquirers but is not actually found in the PLD. In such a situation, the lack of the ‘expected’ feature may be a kind of evidence: indirect negative evidence.

A sufficiently clear and robust characterization of the markedness of parameter-settings may provide a form of indirect negative evidence. If the marked value of a parameter associated with a given feature is that associated with movement, then if there is no evidence for movement in the PLD, the acquirer has indirect evidence that the marked value of the parameter in question does not hold. In other words, evidence for marked features requires direct positive evidence, and indirect negative evidence that the positive setting does not hold arises simply when the direct positive evidence is not available. That this is the case follows from our basic characterization of the asymmetric nature of markedness relations: the presence of the marked feature must be in some way signalled. So if there is no evidence for a marked parameter value, it is not assumed. This in itself is a form of indirect negative evidence.

The feature-counting notion of markedness of parameter values that was introduced in §3.4.2 above is a purely formal one. As such, it differs from other approaches which have tried to relate markedness to **substantive universals**, either directly or indirectly. Chomsky and Halle’s (1968) marking conventions relate the purely formal, feature-counting evaluation metric they propose to substantive phonetic and phonological universals. We also mentioned that Cinque (1999: 128) proposes a series of markedness conventions for the features associated with various functional heads in his analysis of clause structure. His postulations of marked and unmarked values are based on familiar Jakobsonian criteria, as we saw. For example, Cinque assumes that the unmarked value of his postulated $\text{Mood}_{\text{Speech Act}}$ category is ‘declarative’, while the marked value is ‘-declarative’; the unmarked value of $\text{Mod}_{\text{epistemic}}$ is ‘direct evidence’ and the marked value is ‘-direct evidence’, the idea being that in each case the unmarked value is inherently simpler than the marked one.

How do Cinque’s proposals regarding markedness relate to the proposal made above regarding the relation between complexity and markedness? The two notions are quite distinct, in several important respects. The

fundamental difference between the two is that Cinque's proposals regarding markedness relate to the substantive content of features of functional heads, ultimately their notional semantic properties, while what was sketched in §3.4.2 is a purely formal, feature-counting notion associated with a complexity metric. We therefore might want to keep the two kinds of markedness distinct. We could call the complexity-based notion of markedness discussed above formal markedness and Cinque's notion substantive markedness. (This distinction is proposed in Roberts and Roussou (2003: 214), although on slightly different grounds.)

We saw earlier that Chomsky and Halle (1968, Chapter 9) link formal markedness (their feature-counting evaluation metric) to substantive markedness by means of markedness conventions. We might want to contemplate a similar move in the present context. One reason for this is that, as we saw above, Cinque proposes as one criterion of markedness a greater likelihood of morphological expression. This connects to the postulates introduced in the previous section regarding the morphological expression of certain parameter values.

Let us repeat those statements here:

- (34) If (finite) V is marked with person agreement in all simple tenses, this expresses a positive value for the V-to-T parameter.
- (37) If a DP has morphological dative case, then the grammar has abstract Dative Case.
- (41) If T[-finite] has an Accusative feature, then it shows inflectional distinctions marking tense/aspect.

We can note that, directly in the case of (34) and indirectly (by means of the marked way of realizing SpecTP in the absence of a Nominative DP in the case of (37)), the realization of a morphological feature implies the marked value of the parameter ((41) seems to go the other way, though). (34) and (37) suggest that the following general template might hold for the relationship between morphological expression of a parameter and the markedness of that parameter:¹⁸

¹⁸ (50) is deliberately vague in formulation. The expression 'C is associated with a marked parameter value' is formulated so as to allow for (37), where the connection between morphological dative case and markedness, in terms of the complexity of the structure in (39b') above where the relevant DP is Dative, is somewhat indirect (although it is in fact the consequence of the fact that Dative Case is interpretable and therefore unable to check a feature of T – see the discussion in §2.3.2). If it is anywhere near correct, (50) no doubt requires a great deal of refinement.

- (50) If a formal feature of a category C is inflectionally expressed, then C is associated with a marked parameter value.

Although rather vague as it stands, something like (50) could serve as a marking convention linking overt inflectional morphology with the marked values of syntactic parameters, as well as providing a clear general statement of the kind of morphological triggers (or cues, in Lightfoot's (1999) terminology) that are relevant in acquisition and change. It also predicts that the loss of inflectional morphology, at least for certain types of inflection, may perturb the PLD in such a way as to lead to abductive change along the lines we saw in the previous section. Taking (41) into consideration suggests that the implication might go either way, but we nevertheless observe a relation between morphology and the changing and setting of parameters. In the next section we will propose a further marking convention related to cross-categorical harmony in word-order patterns and word-order change.

3.4.6. *Markedness, directionality, and uniformitarianism*

One final very general point regarding markedness concerns the concept of uniformitarianism. We briefly mentioned this concept in §2.4 in our discussion of diachronic aspects of subordination. This idea is formulated by Croft (2003: 233) as follows: '[t]he languages of the past ... are not different in nature from those of the present'. In terms of the principles-and-parameters approach to syntax, we can take this to mean that all languages at all times (in the history of our species) reflect the same basic UG and therefore the same set of parametric options, and that those parametric options have the same markedness properties.

Stated as above, the uniformitarian hypothesis seems very plausible. In fact, one can argue that it is a precondition for applying the principles-and-parameters approach to diachronic questions (see Roberts (2001: 89)).¹⁹

¹⁹ Or indeed any kind of historical linguistics. Interestingly, for most of the history of linguistic thought in the West, uniformitarianism was not assumed, in that it was thought that the three languages of the Holy Scriptures, Latin, Greek, and Hebrew, were not subject to change or decay. See the discussion of Dante's *De vulgari eloquentia* in Law (2003: 190, 230). Clearly, the assumption that Latin and Greek could change was necessary for comparative Indo-European philology to be possible, although the Renaissance recognition of the changeability of Latin did not give rise to the postulation of the Indo-European family (see Law (2003: 260ff.) and Simone (1998: 215) on this).

However, one question we can raise has to do with the transition from unmarked to marked parameter values, an issue we have postponed until the next section (see note 16). We clearly want to allow for the transition to marked parameter values, although we have not yet seen how this may be possible in terms of the general approach outlined above. If we do not allow for the innovation of marked values, then two highly problematic issues arise. First, we predict that all languages are tending towards a steady state, from which they will not be able to escape, where all parameters are fixed to unmarked values. Second, we have to explain where the marked parameter values currently observable in the world's languages came from. So it is highly desirable to have a mechanism for the innovation of marked parameter values.

The question of uniformitarianism arises here, in that if every language were in the maximally steady state we would have a violation of a strong version of this thesis. However, at least a weaker interpretation, of the kind just given in terms of principles and parameters, would allow for the idea that change from marked to unmarked is more regular and frequent than change from unmarked to marked. This would entail that the set of languages in the world would gradually change towards ever less marked systems. On this view, UG and the available parameters do not change, and so uniformitarianism is not violated, but at the same time the range of different sets of options actually instantiated in the world's languages steadily diminishes. In other words, if we think of the set of parameters as defining an abstract space (perhaps a 'state-space' in the terminology of dynamical systems – see §4.3.3) within which grammars can exist, a general move towards more and more unmarked systems implies that ever smaller pockets of the available space are occupied by actually existing systems. Something like this is certainly possible in principle; whether it is actually happening is an empirical question, albeit a rather difficult one to answer with any certainty. At first sight, there appears to be some evidence for something like this from typological studies: Nichols (1992: 250–1), for example, observes that the overall level of structural diversity in (some aspects of) grammatical systems is lower in the Old World than in the New World and the Pacific. She points out that '[t]he high diversity there [in the New World and the Pacific – IGR] can be regarded as a peripheral conservatism in dialect-geographical terms; these areas, secondarily settled, are far enough from the Old World centers of early spread to have escaped the developments that have lowered genetic density and structural diversity in the Old

World' (250). However, determining whether there is a global tendency for reduction in diversity requires knowledge of change at very great time depths, greater than the maximum of 8,000–10,000 years which the traditional method of comparative reconstruction seems to allow. Nichols (1992) addresses this very question, and in fact concludes that 'today's linguistic universals are the linguistic universals of the early prehistory of language' (Nichols 1992: 278). This conclusion strongly favours the uniformitarian view, and the concomitant view that the world's languages are no less evenly spread among the options made available by UG than they were in prehistory. As Nichols states '[t]he only thing that has demonstrably changed is the geographical distribution of diversity' (277). We thus clearly need a mechanism for introducing marked parameter values, as there seems to be an overall equilibrium over time in the grammatical systems attested, as far as we can tell; I will return to this point in the next section.

3.4.7. *Conclusion*

In this section we introduced the concept of markedness and applied it to parameter values, in terms of the definition of complexity given in (22). We also looked at other approaches to the markedness of parameter values, notably the Subset Principle. Further, we briefly considered the relationship between markedness and indirect negative evidence, as defined by Chomsky (1981), and Cinque's (1999) proposals for substantive markedness values associated with functional heads. We considered the relationship between inflectional morphology and syntactic markedness, tentatively suggesting the correlation in (50). Last, the possibility that the world's languages are tending towards ever more unmarked systems was considered and rejected, following Nichols' (1992) conclusions.

In the next section I will try to conclude the general discussion of parameter-setting which has been the theme of this chapter by making a proposal for the form of parameters and considering some of its consequences.

3.5. **Parameter setting and change**

In this section I attempt to synthesize the discussion in the preceding sections, by proposing a general format for parameters and suggesting an

account of how they are set in language acquisition and change. The goal of this exercise is to give a clear view of the issues involved, and to bring together the strands of the discussion in the rest of this chapter, rather than to make new theoretical proposals. The conclusions we reach here will also form the basis of the discussion in much of the remaining two chapters. Accordingly, we first present a general statement of the format for parameters, basing ourselves fairly closely on Dresher's (1999) formulation, as illustrated in (31) above. We then flesh out further the discussion of markedness from the previous section, presenting a further markedness convention (in addition to that presented in (50)), and showing how the concept of **markedness reversal** may play a role in certain types of syntactic change, primarily word-order change. This leads naturally to a discussion of networks of parameters; here we summarize the very interesting proposals in Baker (2001). The final question we look at, although rather briefly since it will be taken up in more detail in §4.3.4, has to do with 'cascades' of parametric change: the extent to which one parameter change may lead to another and how, once again, markedness considerations may play a role.

3.5.1. *A format for parameters*

The first issue concerns a general statement of the exact form of parameters. This is something that we have not broached until now, having contented ourselves with rather informal statements when we introduced the various parameters we have been considering in Chapters 1 and 2. Let us first recapitulate those statements:

- (51) A. Does every finite clause require an overt subject?
 YES: non-null-subject languages (French, English ...).
 NO: null-subject languages (Italian, Spanish, Greek, Japanese ...)
- B. Does V move to T in finite clauses?
 YES: French, Welsh, Italian, Icelandic ...
 NO: English, Swedish, Danish ...
- C. Does the finite verb move to C in finite main clauses?
 YES: German, Dutch, Swedish, Icelandic, Danish, Kashmiri, Romansch ...
 NO: English, French, Italian, Welsh ...
- D. Are (non-inverse) Negative Agree relations found?
 YES: French, Italian, Welsh ...
 NO: English

- E. Does a *wh*-phrase move to the Specifier of an interrogative CP?
YES: English, Italian, Spanish, German, Welsh ...
NO: Chinese, Japanese, Thai, Korean, Turkish, Armenian ...
- F6. For all heads H, does the structural complement of a head H precede or follow H in overt order?
PRECEDE: Malayalam, Turkish, Japanese, Basque ...
FOLLOW: Romance, Celtic ...
- G. Does L allow accusative subjects in SpecTP of a non-finite clause?
YES: English, Latin, Classical Greek, Irish ...
NO: French, Italian ...

Each of these parameters is formulated as a yes/no question, or, in the case of F6, as a disjunctive question (precede vs. follow). As we pointed out in §1.1, one could imagine that the two-year-old mind/brain has some means of interrogating the PLD along these lines. It is now time to try to flesh this rather crude notion out in a more precise fashion.

We can immediately observe that (51B, C, E) have to do with triggering movement, and can thus conclude that the variation is due to the presence or absence of the movement-triggering feature on the head in question (finite T in (51B), finite C in (51C), and interrogative C in (51E)). Furthermore, if word-order variation is to be accounted in terms of movement relations, as suggested in §2.5.4, we may be able to see parameter F6 (or the group of parameters which determine head-complement order for a range of heads) as a case of the presence or absence of a movement-triggering feature. (51A) and (51G) are slightly different in that they concern the type of subject which can appear in SpecTP: whether there can be a null subject in the Specifier of a finite T or an overt Accusative subject in the Specifier of a non-finite T. Assuming that the possibility of a null subject of finite T is connected to 'rich' agreement, a property we can associate with T, then both of these parameters have to do with the nature of T's features and therefore what kinds of elements T may Agree with. Finally, (51D) concerns the possibility of a particular feature entering a particular type of Agree relation.

So we can draw two conclusions. First, the parameters all concern formal operations of the syntactic system: Agree and Move. They do not seem to relate directly to morphological, phonological, or semantic properties of language. Second, we can see that all of the parameters relate to the features associated with heads; in fact they all involve the features of functional heads, except perhaps for some cases of F6. Again, the features

in question are all formal features, i.e. they are features which play a role in determining the application of formal operations such as Move and Agree. So, all the parameters we have looked at involve the formal-feature specification of heads, principally functional heads; this corresponds exactly to what is proposed in Chomsky (1995: 6).

These observations make possible a general statement of the form of parameters. Following Dresher's (1999) approach, as illustrated in (31) above, we present each parameter along with a statement of its default value and its cue. We will continue to present parameters as binary in nature. On the basis of the discussion of markedness and complexity in the preceding section, we assume that the default value must involve a smaller number of features than the marked value (see (22)). Also on the basis of the discussion in the preceding section, we take it that the cue may be a morphological property. The other obvious cue is word order itself.

The general format for parameters will thus look like this:

- (52) a. *Parameter*: A (functional) head H {has/does not have} feature F (in a given formal relation).²⁰
 b. *Default*: F is absent.
 c. *Cue/expression*: properties of inflectional morphology and linear order of elements.²¹

²⁰ Given the nature of the syntactic operations postulated in recent minimalism, we really only have four options for (52a), and these are implicationally related, as follows:

- i. does H have a feature triggering Agree?
- ii. if so, does H have an EPP feature?

If we distinguish head-movement and XP-movement, then we have two further options:

- iii. if (ii), does H require pied-piping of the Goal?
- iv. if so, how large a category is pied-piped?

This last option was implicit in our discussion of 'massive movement' in §2.5. For more technical and empirical details, see Richards and Biberauer (2005); Biberauer and Richards (2006); Biberauer and Roberts (2005a). I will not pursue the options in (i–iv) systematically here, although the statement of the parameters in (54) is not incompatible with them. We will encounter pied-piping again in §4.1.4, when we look more closely at the nature of formal optionality.

²¹ See §3.2 above for a discussion of the similarities and differences between Clark and Roberts' (1993) notion of P-expression and the Lightfoot/Dresher notion of cue.

Introducing parameters in §1.1, I pointed out that they have four important properties. These can be summarized as follows:

- (53)
- a. Parameter values must be able to be set on the basis of rather salient elements of the PLD.
 - b. Parameter values must be set: not deciding is not an option.
 - c. Parameters may be determined by ‘gaps’ in UG principles.
 - d. Parameters are binary.

How does the schema in (52) capture the properties of parameters as listed in (53)? Let us consider (53a–d) one by one. (53a) clearly relates to the cuing or expression of parameters. The linear order of constituents and inflectional morphology are both salient features of the PLD, and are both things that acquirers appear to be sensitive to, given that they are able to set word-order parameters very early (as we saw in §3.1) and they acquire the morphological properties of verbs, including agreement and finiteness marking, equally early. (This is shown in detail by Guasti (2002: 120ff.)) So the schema in (52) can clearly capture this property of parameters.

Example (53b) relates partly to the default clause in (52), in that we can assume that in the absence of a clear expression of the value of a given parameter (i.e. if all the relevant PLD is weakly P-ambiguous in the sense defined in (21c)), the default option is always taken. A further point which comes up here concerns the relations among parameters. We saw in §1.5.1 that there is a further parameter distinguishing among languages with the positive value for parameter E (i.e. those with overt *wh*-movement) determining whether just one *wh*-phrase is moved to an interrogative C or whether all available *wh*-phrases must be moved. Naturally, this further parameter is not relevant in systems where parameter E has the negative value. We take it that this parameter must take on the default value in this kind of case. What is at issue here is the question of the implicational relations amongst features, a point I will return to below.

Example (53c) can be reconciled with (52) if we make the obvious inference from (52) that, to some extent, the feature make-up of functional (and perhaps some other) heads is underspecified by UG. It may be that UG only requires a very minimal feature specification for functional heads: just enough to distinguish what Chomsky (2000; 2001) refers to as the ‘core functional categories’: C, T and *v*. Further specification may be entirely a matter of parametric variation. (Giorgi and Pianesi (1997, §1.4) make a proposal similar to this.) Indeed, to the extent that the formal features of functional categories are primarily relevant for the internal workings of

syntax, and that, in the context of the Minimalist Program, these internal workings of syntax are as elementary as possible, it seems very reasonable to think that UG imposes no particular further requirements on the feature make-up of functional categories. So we see that (53c) can be captured by the format in (52). This gives us a way to understand why parameters exist at all, which we hinted at in Chapter 1: they simply force a consistent choice where UG leaves things open, i.e. an individual system cannot have gaps and does not tolerate randomness. Every underspecified point must be ‘filled in’ in a consistent way. (The consistency might in fact be created by the learning device; I will briefly take this point up again in §5.1.) Finally, (53d) is built in to the statement in (52a).

Since (52) seems to capture the important properties of parameters as listed in (53), we will take it to be a general format for the statement of parameters. More specifically, the parameters in (51) can now be reformulated as follows:

- (54) A. *Null subjects*
- a. *Parameter*: Finite T {has/does not have} sufficient specification of agreement features φ to bear the subject thematic role/Agree with *pro* in SpecTP.²²
 - b. *Default*: φ is absent.
 - c. *Cue/expression*: ‘rich’ agreement morphology on T- and/or V-elements.
- B. *V-to-T movement*
- a. *Parameter*: Finite T {has/does not have} an EPP feature which attracts V.
 - b. *Default*: EPP is absent.
 - c. *Cue/expression*: (finite) V is marked with person agreement in all simple tenses.
- C. *Verb second*
- a. *Parameter*: Finite, root C {has/does not have} an EPP feature which attracts T.
 - b. *Default*: EPP is absent.
 - c. *Cue/expression*: consistent XP V order in the left periphery of CP (see Lightfoot (1999: 153)).
- D. *Negative concord*
- a. *Parameter*: non-inverse Agree_{Neg} relations.
 - b. *Default*: such relations are absent.
 - c. *Cue/expression*: clausal negation which either can or must be uninterpretable.

²² Recall that in our discussion of the null-subject parameter in §1.1, we did not decide between these two analyses of null subjects.

E. *Wh-movement*

- a. *Parameter*: [+wh] C {has/does not have} an EPP feature triggering movement of a wh-phrase to its Specifier.
- b. *Default*: EPP is absent.
- c. *Cue/expression*: ‘displaced’ wh-phrases, wh-marking on D.

F. *The head parameter(s)*

- a. *Parameter*: a head H {has/does not have} an EPP feature triggering movement of its complement to its specifier.
- b. *Default*: EPP is absent.
- c. *Cue/expression*: overt complement > head orders.

G. *Accusative + infinitive*

- a. *Parameter*: non-finite T {has/does not have} features which Agree with a DP in its specifier.
- b. *Default*: such features are absent.
- c. *Cue/expression*: overt Accusative subjects of infinitives.

So we have a general format for parameters, which seems to have the right kinds of properties, and we are able to reformulate the parameters which we have been interested in by using this format. At the very least, this is a useful exercise, but the combination of the requirement to state each parameter in terms of formal features and to state both the default value and the cue has clear implications for both language acquisition and language change. In essence, our expectation is that, if the cue is not sufficiently robustly attested in the PLD, the parameter will revert to its default value. Given the discussion in the preceding sections, we can see that this has clear implications for both acquisition and change. Thus the exercise of formulating parameters along these lines is one which amounts to making empirical predictions in these two domains.

3.5.2. *A markedness convention for syntax*

In terms of (54), it is easy to see how a parameter changes from a marked to a default value, and it is easy to see how the default values are related to the general simplicity metric in (22), since in each case some feature (or, in the case of (54D), a relation) which is present in the marked state is absent in the default state. But, as we discussed at the end of previous section, we must allow for change in the opposite direction too. One way to do this is by considering how markedness considerations may relate to *systems* of parameters, or perhaps subsystems of related parameters, rather than to

individual parameters. Hence, rather as in the case of distinctive features as discussed by Chomsky and Halle (1968, Chapter 9), it may be that the markedness of a particular parameter will depend on the values assumed by other parameters in a given system. Let us explore this idea further, and consider an illustrative possibility.

We saw in §1.5.1 and §2.5 that the head parameter (54F) is rather problematic as stated. If it were a single parameter, it would predict a spectacular clustering of properties, which is not actually attested in the majority of languages. As we mentioned in the discussion of word-order correlations in §1.6.1, Dryer (1992) shows that a minority of the languages in his sample actually conform to the predictions of this putative parameter across the whole range of head–complement relations. The majority of languages diverge at least in some respects. We suggested in Chapter 2 that (54F) should in fact be broken down into a series of related parameters relating to each head–complement pair. However, without some further statement, all predictions regarding word-order correlations are thereby lost. The preference for ‘harmonic’ ordering seems to derive from an overriding tendency for independent parameters to conspire to produce a certain type of grammar. To capture this, we tentatively suggested that a restatement of J. Hawkins’ (1983) generalization regarding cross-categorial harmony is needed, along the following lines (repeated from (84) of §2.5):

- (55) There is a preference for the EPP feature of a functional head F to generalize to other functional heads G, H ...

Now it is time to relate (55) to the ideas about markedness we have been developing. We can think of (55) as an approximation to a markedness convention of the type proposed for phonology by Chomsky and Halle (1968).

To take a specific example, suppose, following Kayne (1994) and the discussion in §2.5.4, that VO is the universal underlying order and that OV orders derive from the combination of V-to-v raising and remnant VP-fronting to SpecvP, as illustrated in (56):

- (56) [_{VP} [_{VP} O (V)] v+V (VP)]

In terms of (54F), v has a marked property here. Following Chomsky and Halle’s notation, let’s write this as the *mEPP* value for v.²³ In rigidly head-final languages like Malayalam (see §1.5.1), many, perhaps all, functional

²³ Presumably v actually has two EPP features, since it attracts both V and VP. Here I am only concerned with the one which attracts VP.

heads will have at least one EPP feature in this way. Such systems will therefore emerge as very marked indeed, in terms of what we have said so far, and yet they are more common than ‘mixed’ types like Latin, German, etc., which would be less marked on this approach.

It is here that markedness conventions and the concept of the markedness of a whole system, or subsystem, of parameters comes in. Let us postulate, for concreteness, the following convention:

- (57) For a class of heads H , μEPP for $H_{\text{uF}} \neq v \rightarrow \left\{ \begin{array}{l} [+EPP] / v_{[+EPP]} \\ [-EPP] \end{array} \right\}$

What (57) says is that the unmarked value of the EPP feature for some head of a particular type with an uninterpretable feature (i.e. a Probe, capable in principle of triggering movement) is $[+EPP]$, i.e. the presence of an EPP feature, just where v has an EPP feature, i.e. in an OV system. (Here the EPP feature is understood to refer to attraction of VP rather than V; see note 23.) This convention would replace the default statement associated with the head parameter in (54F, b). This has the effect that, for all head–complement pairs which are subject to word-order variation, head-final is the unmarked order in an OV system, and head-initial in a VO system. In these terms, rigidly head-final languages are relatively unmarked, as of course are rigidly head-initial languages, while ‘mixed’ languages are relatively marked (and one can in principle quantify exactly how marked different types of mixed systems would be). Furthermore, Dryer’s observation that VO vs. OV order is the basic determinant of ordering among other head–complement pairs is directly captured by (57). (See again the discussion of Dryer’s results in §1.5.1.) What remains unclear, however, is how to specify the class of heads (57) refers to.

A possible disadvantage of (57) is that it appears to disconnect markedness from the simplicity metric in (22), in that we are now claiming that systems where the EPP feature is present on all possible heads are relatively unmarked. Hence a simple feature-counting approach to simplicity and thence to markedness no longer suffices. However, we can think that the simplicity metric itself derives from a more general notion of the conservatism of the learner, in that the learner will strive to assign the simplest representation or derivation possible to the PLD it is exposed to. In these terms, we can understand a markedness convention like (57) in terms of the conservatism of the learner, assuming that another conservative aspect of

the learner would be to exploit pieces of, perhaps marked, input to the full. So we could entertain something like the following:

(58) Generalization of the input:

If acquirers assign a marked value to H, they will assign the same value to all comparable heads.

Example (57) can naturally be understood in terms of (58), and both (57) and (22) can be seen as different aspects of the overall conservatism of the learner, which is essentially trying to set parameters in the most efficient way possible. The Subset Principle, as discussed in §3.4.3 above, can also be seen in this light: one aspect of the learner's conservatism is to avoid superset traps.

3.5.3. *From unmarked to marked*

A markedness convention like (57) also gives us a way of seeing how individual parameters may change from an unmarked to a marked value. All heads which are capable of bearing EPP features have the inherently unmarked property of not bearing this feature, but, where *v* has an EPP feature the opposite is true. Thus, if *v* acquires an EPP feature, a markedness reversal takes place for all the other heads in the system, and this creates pressure, ultimately due to (58) as a property of the learner, in the direction of acquiring the [+EPP] value for all other heads. Of course, this does not answer the question of how *v* might acquire an EPP feature. One might object to this approach along the same lines as Song's (2001: 304) objection to the Lehmann–Vennemann approach to word-order change discussed in §2.5.2. The preference for markedness-induced harmony must be weak enough to permit 'incongruous' word orders to arise in otherwise consistent grammars but strong enough to cause 'endogenous optimization' in Kiparsky's (1996: 150ff.) sense, i.e. harmonization of the relevant attraction properties of other functional heads. Kiparsky (1996: 153) defends his position as follows:

We can therefore legitimately posit a universal preference which is not universally instantiated, provided that we specify the other factors that allow (or force) it to be subverted. In principle, they might be either intersecting structural or functional constraints, or historical processes. Motivating the latter would, in the case at hand, amount to demonstrating a natural origin for OV syntax.

Kiparsky goes on to suggest that OV syntax could arise from a system in which objects (and perhaps other arguments) are in apposition to pronouns,

and hence frequently left-dislocated for reasons to do with information structure (topicalization, focalization, etc.). If the pronouns become agreement markers and/or disappear, OV syntax may emerge. Although speculative, this scenario gives an indication of how a dispreferred, possibly incongruous, order might arise through a separate, natural kind of change.

We can readily rephrase Kiparsky's speculation using the formal notions adopted here. In these terms, the question of the origin of OV orders becomes the question of how *v* could acquire an EPP feature. The most likely scenario for *v* acquiring an EPP feature is that whereby optional, discourse-driven object-movement becomes obligatory through the loss of the discourse effect, as Kiparsky suggests. This can be fairly naturally stated in terms of some of Chomsky's recent assumptions, since he allows for optional movement triggers (i.e. EPP features) as long as their presence has an effect on output, i.e. creates some kind of discourse effect (Chomsky 2001: 34). At an earlier stage *v*'s optional EPP feature would be dissociated from its uninterpretable φ -features (which might, following Kiparsky's suggestion, Agree with a resumptive pronoun in object position), but later the two sets of features would coalesce and the EPP feature would thereby be obligatorily associated with Agree in φ -features with the object, giving rise to object-movement (or possibly VP-pied-piping, as described in §2.5.4).

That the discourse effect is associated with a complication of structure through the imposition of an extra EPP feature gives a formal expression of the traditional intuition that the drive for expressivity is a factor in language change alongside the drive for simplicity (see Martinet (1955) for similar ideas in the context of sound change), and that, in the long run, these two forces create an overall equilibrium. This may be what prevents languages from developing the maximally unmarked steady state. Nichols' (1992) evidence that the overall degree of diversity in the world's languages has not changed since prehistory supports the idea that marked structures must be able to be innovated; in the terms just described, the tension between expressivity and simplicity balances out over the very long term, and there is thus no net increase or decrease in the markedness of the systems that are attested at any historical moment.

Expressivity may cause EPP features to be introduced, while simplicity causes them to be eliminated. Presumably, a constraint like (58) causing generalization of the input causes them to become obligatory. Furthermore, it is very likely that the 'coalescence' of EPP and φ -features alluded to above is driven by the preference for simplicity. We can thus envisage a

sequence of changes, starting from an optional EPP feature, to obligatory EPP combined with Agree, to simply Agree. For *v*, this would give rise to a sequence of changes from optional OV associated with a discourse effect, to obligatory OV associated with case marking and/or agreement, to VO. As already mentioned, this is fully consistent with the proposals in Kiparsky (1996).²⁴ Clearly, these speculations require a great deal more work before they can really be considered as established hypotheses, but we can see in principle at least how relatively marked systems might be innovated, and hence avoid the problems with positing marked and default parameter values which we noted in §3.4.6 above.

3.5.4. *Networks of parameters*

Looking at the ways in which parameters may interact in change and in determining the markedness of a whole system leads naturally to the question of networks of parameters. We already saw a very simple example of how one parameter may determine the value of another one in our brief discussion of the relation between the parameter determining multiple wh-movement and parameter (54E). We can think of this relation as a kind of intrinsic ordering, in that the multiple wh-movement parameter depends on the value of the superordinate parameter (54E): if (54E) has the negative value, then the multiple wh-movement parameter can only have the default value; it is effectively ‘switched off’ as an independent parameter.

A natural question to ask is how far this kind of intrinsic ordering among parameters can be taken. It is clearly a desirable feature of a parametric system, as it automatically ensures certain empirical predictions. If P_1 is superordinate to P_2 in the sense just defined, then P_1 must be acquired before P_2 and a diachronic change in the value of P_1 will potentially affect the value of P_2 , but not vice versa.

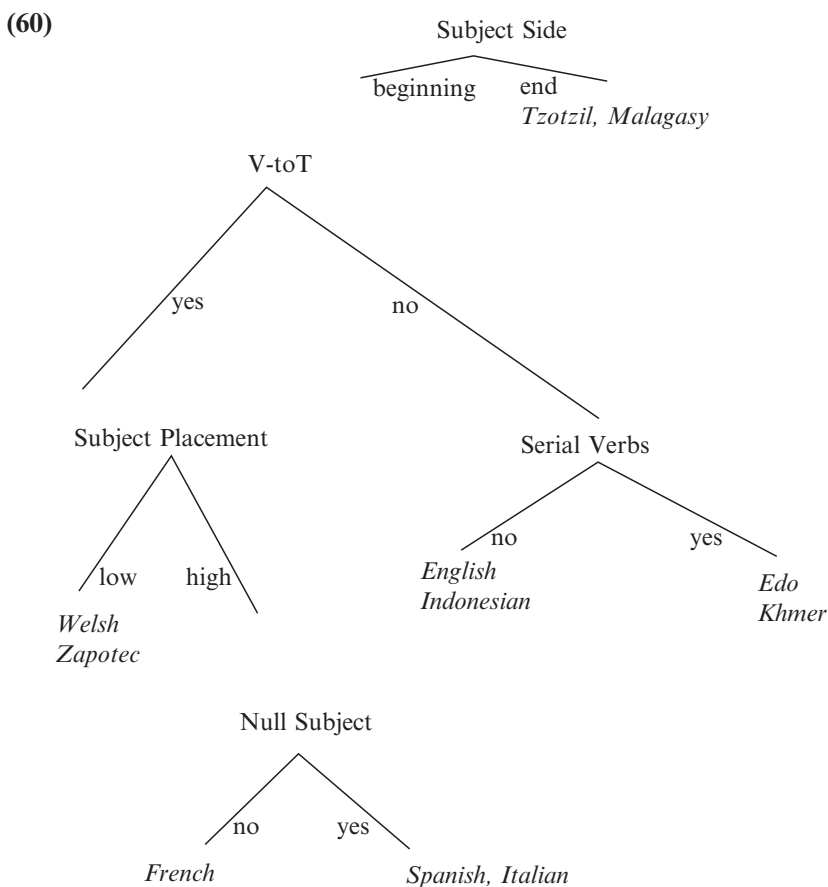
Baker (2001: 163) formulates exactly this notion of intrinsic ordering among parameters as follows:

²⁴ In terms of Chomsky’s (2005c) proposal that there may be a further variety of movement which is entirely separate from Agree, triggered by the Edge Feature EF (see note 17), we might replace the optional EPP feature at the first stage of the cycle with an EF feature. In that case, *v* changes through all the formal options the current theory makes available.

- (59) Parameter X ranks higher than parameter Y if Y produces a difference in one type of language defined by X, but not in the other.

Keeping to our rather simple example involving wh-movement for the purpose of illustration, (54E) would be parameter X in Baker's formulation, and the multiple wh-movement parameter would be parameter Y, since the latter produces a difference only in those languages with overt wh-movement and not in those without.

Baker develops what he calls a 'periodic table' for parameters using the notion of ranking (what we have been calling intrinsic ordering) in (59). A subpart of this, which involves some of the parameters in (54), is given in (60) (Baker's Figure 6.4, 183, presents the full system he proposes):



V-to-T is the familiar parameter (54B), and Null Subject is our (54A). We have seen the Subject Placement parameter, notably in §1.2.1; this parameter determines whether the subject raises to SpecTP. Baker's Subject Side parameter determines, roughly, whether the subject appears at the beginning of the clause, or at the end. Tzotzil and Malagasy are VOS languages, in which the canonical position of the subject is final, while all the other languages listed here are SVO or VSO, depending on the value of the Subject Placement parameter; the head parameter is superordinate to all the parameters in (60).²⁵ The Serial Verb parameter determines how many verbs a single VP (or perhaps vP) can contain (Baker 2001: 141). English doesn't allow more than one (main) verb per VP/vP, but many languages do, for example Edo (a Niger-Congo language spoken in Nigeria) does:

- (61) Òzò ghá lè èvbàré khiè'n.
 Ozo will cook food sell
 'Ozo will cook the food and sell it.'
 (Baker 2001: 140)

What (60) actually states is a series of intrinsic ordering relations among parameters: the Subject Side parameter is superordinate to the others given here in that only the 'beginning' value for this parameter allows a choice in the V-to-T parameter, since the clause-final position of the subject means that the position of V in T or lower would not affect word order. Similarly, only the positive value of V-to-T allows an option regarding Subject Placement, since the negative value will result in SVO order whether or not V moves. Only a negative value of V-to-T is compatible with the option of serial verbs, since if there are two verbs in vP/VP it is impossible for both to raise to a single T-position.

A hierarchy of parameters of the kind in (60) makes interesting predictions regarding typology, acquisition, and change. Regarding typology, it

²⁵ It has recently been proposed that VOS order should be derived by raising VP (excluding the subject, which is taken to be merged in SpecvP) to SpecTP. This was first proposed by Massam and Smallwood (1997); see also Massam (2000; 2005); Rackowski and Travis (2000); Chung (2005). This analysis would be consistent with Baker's proposals as given in (60), as VP-fronting to SpecTP would arguably 'bleed' both V-to-T movement (since the verb must remain in the fronted VP for the VOS word order to result from VP-fronting) and subject raising, if VP-movement satisfies T's EPP-feature, as suggested in the references just given.

predicts a series of implicational universals: if a language has serial verbs, it is SVO (VSO depends on the positive value of V-to-T, and VOS on the 'end' value for Subject Side; recall that OV languages are determined by the higher-order head parameter); if a language has null subjects, it has V-to-T; if a language is VSO, it does not have serial verbs. Clearly, all of these predictions are testable (and some of them are false: for example, (60) predicts that if a language is VSO it does not have null subjects, but Welsh, Irish, and Classical Arabic are all VSO null-subject languages; this point is also made by Newmeyer (2004: 201; 2005: 86)).

In the domain of language acquisition, (60) predicts what Dresher, following Lightfoot (1989), calls a 'learning path'. (Baker (2001: 192–6) also makes this point.) The setting of a superordinate parameter will determine whether or not there is an option to set a subordinate parameter. For example, choosing the 'end' value of Subject Side pre-empts the setting of any of the other parameters in (60). In terms of what we said earlier regarding parameter interactions and the schema for parameters in (52), we could take this to mean that all the subordinate parameters automatically take on the default value. (This would be the case because all the PLD would be weakly P-ambiguous in the sense of (21).) We thus predict that acquirers of Italian follow a learning path starting from Subject Side (beginning), and going on to V-to-T (yes), Subject Placement (high) and Null Subject (yes). Acquirers of English, on the other hand, set V-to-T to the negative value and then the Serial Verb parameter to the negative value. Again, the predictions for language acquisition are clear in principle. However, once again, the evidence for very early parameter-setting discussed in §3.1 makes it difficult to test these in practice (Baker's conclusion is slightly more optimistic than this, however).

Finally, (60) makes interesting predictions about relations among parametric changes. For example, if a language loses V-to-T movement, then it simultaneously loses the possibility of having VSO order or null subjects, but may go on to develop serial verbs. The history of English since the loss of V-to-T in the Early Modern period is consistent with this, but only in a rather unrevealing way, since serial verbs have not in fact developed. English- and Romance-based creoles, on the other hand, support this, in that they tend to lack V-to-T and (argumental) null subjects, and to have SVO order and serial verbs. (See Muysken (1988); the papers in DeGraff (1999); Nicholis (2004); and §5.3.2 on the syntactic properties of creoles.)

There are two obvious objections one can make to (60). First, it is empirically incorrect, in that certain pairs of parameters are set in the wrong relation with one another. We mentioned an example of this above: the incorrect prediction that there are no null-subject VSO languages. Another incorrect prediction is that there are no null-subject languages which have serial verbs: many East Asian languages, including Chinese, Thai, and Vietnamese, appear to show both of these properties. However, such difficulties can be dealt with quite easily; the hierarchy simply needs to be appropriately revised. The second difficulty is perhaps more serious: Baker allows for the possibility that a given pair of parameters may be logically independent, and in fact discusses (184ff.) several well-established parameters which do not appear to fit into his hierarchy, notably the *wh*-movement parameter (54E). But of course if too many pairs of parameters are independent from one another the hierarchy may start to lose its clear structure. The greatest difficulty would arise if one could show that, for a triad of parameters P_1, P_2, P_3 , P_1 is superordinate to P_2 , P_2 is superordinate to P_3 , but that P_3 is superordinate to P_1 ; this would create a kind of ordering paradox, since the relation ‘ P_n is superordinate/subordinate to P_m ’ is, one assumes, logically transitive. It is not clear whether a case like this actually exists. If it does, then the concept of parameter hierarchy would have to be abandoned in favour of a looser notion of network, and presumably some of the wide-ranging predictions that a hierarchy like (60) makes would be lost. As things stand, though, Baker’s proposals, or some variant of this hierarchy, are of great interest especially for establishing connections between acquisition and change, and have not been shown to be unworkable.²⁶

A final point regarding parameter interactions concerns the possibility of ‘cascades’ of changes: a situation where an initial parameter change perturbs a system in such a way that a whole series of changes follows, perhaps over many centuries, creating the appearance of typological drift.

²⁶ Newmeyer (2004; 2005) argues against Baker’s parameter hierarchy. But his main critique is really the same as the first point just made: some of the parameters may be placed in the wrong relationship to one another. As pointed out by Roberts and Holmberg (2005), this is not really a criticism of the concept of a parameter hierarchy (still less of the concept of parameter itself), but rather of Baker’s specific implementation of it. Roberts and Holmberg further take issue with a number of Newmeyer’s criticisms of the principles-and-parameters approach to comparative syntax.

As we noted earlier, Longobardi (2001: 278) mentions this possibility in his discussion of Inertia, pointing out that syntactic change can be the consequence ‘recursively, of other syntactic changes’. This possibility undoubtedly exists, and may be behind the observations of typological drift that have been made. I will defer detailed discussion of this to §4.3.4, where I explore this idea by looking at a series of changes which took place in the history of English between roughly 1400 and 1700. For the moment, the only relevant point is that the intermediate grammars during the sequence of changes must all be relatively highly marked, and therefore prone to change. Clearly, typologically ‘mixed’ systems will be of this kind, given the postulated markedness convention in (57).

3.5.5. *Conclusion*

In this section, we have attempted to consolidate the discussion in the earlier sections of the chapter, and to some degree that in the earlier chapters as well, by considering the format for parameters and the various ways in which parameters may interact, giving rise to networks or hierarchies, as well as the concept of markedness of an entire system. As we saw, markedness of a system may override the markedness specification of an individual parameter.

3.6. **Conclusion to Chapter 3**

This chapter has attempted to consolidate the ideas which were introduced in the first two chapters. There we first tried to demonstrate the utility of the notion of parameter of UG for analysing syntactic change (Chapter 1) and for giving a (near-) unified account of different types of change (Chapter 2). Here, we tried to show how parameter change can be seen as driven by language acquisition. The essential notion is that of the conservatism of the learning device, which always attempts to set parameters on the basis of the greatest computational efficiency. This has at least two consequences that we have seen: a strong tendency to favour relatively simple representations or derivations, which we stated as (22); and a tendency to generalize the input, which we formulated as (58), underlying the markedness

convention in (57). Both of these properties motivate a formulation of the default values of parameters (although (57) concerns the markedness of systems, and as such may create a markedness preference which overrides the markedness value of a single parameter). This contributed to our general statement of the format of parameters in (52), inspired by the proposals in Dresher (1999), whereby each parameter consists of a formal statement (itself highly restricted by the impoverished mechanisms of minimalist syntax – see note 20), a statement of the (defeasible) default, and a statement of the cue or expression of the parameter. This format seems to have the right properties and is useful for looking at both acquisition and change.

In seeking to relate parametric change to language acquisition, we undertook a survey of recent work on the acquisition of syntax in §3.1. Here we encountered the Very Early Parameter Setting observation, which to some extent hampers establishing a straightforward relationship between acquisition and change, although it does not preclude such a relationship. §3.2 discussed the logical problem of language change, which led us to the formulation of the simplicity metric in (22). The subject of §3.3 was the changing trigger. Given Inertia, i.e. the idea that syntactic change must be caused (to paraphrase Longobardi (2001: 278)), we considered how contact and morphological erosion may induce change. §3.4 dealt with markedness at some length. We suggested that (22), along with various markedness conventions in the pattern of Chomsky and Halle (1968, Chapter 9), should form the basis of markedness. The Subset Principle may also be relevant if parameter systems allow formal optionality. Finally, in §3.5 we arrived at our formulation of parameters in (52) and considered its implications for networks and hierarchies of parameters, paying particular attention to the proposals in Baker (2001). We also made a suggestion for how marked properties may be innovated, at least in the case of EPP features. This suggestion seems to capture the old idea that much of language change is caused by a tension between a drive for simplicity and a drive for expressivity.

In the next two chapters, we look at the consequences of the general view of syntactic change that we have outlined over the preceding chapters. We begin, in Chapter 4, by looking at the dynamic aspect of syntactic change – and considering how it might be handled in the terms described here. Chapter 5 focuses on questions connected to contact, substratum effects, and creoles.

Further reading

Principles-and-parameters theory

Baker (2001) is an excellent introduction to the principles-and-parameters conception of UG. Baker pursues a sustained analogy between contemporary comparative syntax and nineteenth-century chemistry, which culminates in a ‘periodic table’ of parameters, part of which is reproduced in (60). Baker observes that any analogue to the quantum-theoretic explanation of why the periodic table has the properties it has is far off. **Newmeyer (2004; 2005)** argues at length that the principles-and-parameters approach to comparative syntax has failed, and that variation across grammatical systems should be handled in terms of performance systems of various kinds. **Roberts and Holmberg (2005)** is a reply to Newmeyer, arguing that the principles-and-parameters approach is a valid and useful approach to comparative syntax.

Learnability and markedness

Lasnik (1983) is an early discussion of learnability in relation to principles-and-parameters theory. **Lightfoot (1989)** is the first statement of the degree-0 learnability idea, developed at much greater length in Lightfoot (1991). **Fodor (1998)** proposes a learnability theory for syntax. **Roberts (2001)** looks at the relation between syntactic change and learnability, proposing a version of the simplicity-based approach to markedness summarized in this chapter. **Dresher and Kaye (1990)** is the initial proposal for cue-based learning of phonological parameters, later developed in Dresher (1999). **Berwick (1985)** first put forward the Subset Principle as a natural learnability-driven constraint on the language-acquisition process. **Manzini and Wexler (1987)** offer an account of parametric variation involving long-distance reflexives, which makes explicit reference to the subset relations among the languages produced by grammatical systems defined by the different parameter-settings proposed. This represents a further case where the Subset Principle may be relevant for understanding the relations among parameter values, and perhaps as a basis for a theory of markedness. **Niyogi and Berwick (1995)** is a pioneering study of how syntactic

change can be mathematically modelled. **Battistella (1996)** is an introduction to and historical overview of the concept of markedness, with particular reference to syntax.

Language acquisition

Hyams (1986) is the ground-breaking study of the acquisition of syntax using the principles-and-parameters approach, in which the phenomenon of early null subjects is first described. Hyams analysed these as null subjects of the Italian type, an idea she has abandoned in subsequent work (see **Hyams (1992)**; **Hyams (1996)**; **Hyams and Wexler (1993)**). **Radford (1990)** was the first to generalize Hyams' (1986) account, and argue that English children, at least, go through a stage of acquisition in which no functional categories are available at all. This work led directly to the postulation of root infinitives and early null subjects. **Pierce (1992)** is a pioneering study of Early French, in which it is shown that at the root-infinitive stage, the infinitival form of the verb does not raise to T while the optional finite form does. **Poeppl and Wexler (1993)** is an important study of Early German, in which they argue for a root-infinitive stage in that language, and that infinitive verbs do not undergo the verb-second operation (i.e. they do not move to C). **Rizzi (1994)** is an influential study of root infinitives, in which it is argued that these derive from the possibility of 'clausal truncation', i.e. realising a clause as a VP only, at a stage of acquisition in which the language faculty is not fully mature. **Rizzi (2000)** proposes something similar for 'diary-drop'. **Clahsen, Kursawe, and Penke (1995)**; **Clahsen and Penke (1992)**; and **Clahsen and Smolka (1985)** are all studies of Early German, in which it is shown that the complex adult verb-movement system develops according to a series of well-defined stages. **Guasti (1996; 2000)**; **Haegeman (1995a)**; **Haegeman (1995b)**; **Hamann and Plunkett (1998)**; and **Hoekstra and Hyams (1998)** are all studies of the early stages of the acquisition of various Romance and Germanic languages from the perspective of principles-and-parameters theory. Hoekstra and Hyams' article is noteworthy for advocating that Early Null Subjects of the kind found in non-null-subject languages such as English and other Germanic languages are not to be equated with those found in null-subject languages such as Italian. **Wexler (1992; 1994; 1999)** provides overviews

and summaries of much of the work in this field, as well as developing more general ideas, notably the Very Early Parameter Setting observation discussed in §3.1. **Friedemann and Rizzi (2000)** is a collection of important articles on the acquisition of the syntax of a range of Germanic and Romance languages. **Brown (1970)** is an early and very influential study of the first-language acquisition of English, while **Jakobson (1941)** is, among other things, a ground-breaking study of language acquisition and language disorders, in which the concept of markedness plays an important role. **Ernst (1985)** is an extremely detailed study of Héroard's journal, in which the speech of the young dauphin was recorded over a period of several years. Héroard's journal is a unique document, of potentially great interest for language acquisition and language change, as well as providing a valuable record of the nature of spoken French in the early seventeenth century. **DeGraff (1999)** is a highly original collection of articles dealing with creolization, language acquisition, and language change. The Introduction and Epilogues are extremely useful and thought-provoking. This collection represents a unique attempt to bring together these areas, which have often been studied somewhat in isolation from one another.

The null-subject parameter

Huang (1984; 1989) develops a 'generalized-control' approach to null subjects in Chinese and Italian, covering also the distribution of the null subject of non-finite clauses in languages such as English (conventionally known as PRO in government-binding theory). **Nicholis (2004)** is a detailed study of the status of the cross-linguistic predictions made by the version of the null-subject parameter put forward in Rizzi (1982), and given in (28) of §1.2.1. He concludes that the correlations hold up fairly well across a wide range of languages, but that the distribution of expletive null subjects in creoles is problematic (see §5.3.2, on these). **Holmberg (2005)** is a recent and very original paper on null subjects, arguing, on the basis of the fact that Finnish has an overt expletive subject which appears to be in complementary distribution with a referential null subject, that null subjects are structurally pronouns.

Other works on syntactic theory

Giorgi and Pianesi (1997) propose a general theory of the syntax of temporal relations, and an analysis of the temporal systems of Italian, English, and Latin. A facet of their approach is the idea that functional heads are structurally present only when needed in order to bear certain features. This view differs notably from that put forward by Cinque (1999). **Massam (2005)** and **Massam and Smallwood (1997)** are analyses of Polynesian languages with VSO and VOS orders in which the central idea is that the V-initial orders derive from VP-fronting, possibly of a remnant VP. **Chomsky (2005c)** is, at the time of writing, the most recent statement of certain technical aspects of minimalism. This paper pays particular attention to the ‘A’-system’, i.e. wh-movement, topicalization, and focalization, all movements to the Specifier(s) (or ‘edge’) of CP. It is proposed that these movements are triggered by the E(dge) F(eature), a feature that has no connection with the Agree system. **Richards and Biberauer (2005)** develop an analysis of the distribution of what have often been seen as overt and null expletives in Germanic (see the discussion of null expletives in §1.2.1) which makes use of the twin notions of ‘massive movement’ of vP to SpecTP (as briefly described in §2.5.4) and the optionality of pied-piping operations. **Biberauer and Richards (2006)** make a very interesting and well-argued case for formal optionality in syntax, arguing in particular that this is a natural outcome of the kind of minimalist syntax proposed in Chomsky (2000; 2001). We will look at some of their proposals in more detail in §4.1.4.

Historical and typological syntax

King (2000) is a detailed and very interesting study of Prince Edward Island French. In addition to arguing convincingly that Preposition-stranding in this variety is the result of extensive borrowing of English prepositions, as summarized in §3.3 above, King looks at the syntactic consequences of the borrowing of the particle *back* and the wh-elements *whoever*, *whichever*, etc., into this variety of French. **Keenan (2002)** is a very detailed, original and interesting study of the development of English reflexives, arguing convincingly that they were originally emphatic forms. It is here that the

Inertia Principle is proposed for the first time. **Kroch *et al.* (1997)** is a detailed study of the loss of V2 in the history of English, arguing that this change was driven by contact between Northern and Southern dialects of ME. **Croft (2000)** puts forward a general account of syntactic change in functional-typological terms, hence differing in its basic assumptions for what is being put forward here. **Nichols (1992)** put forward a number of important innovations in language typology. The distinction between ‘head-marking’ and ‘dependent-marking’ was first made here. This is the distinction between a system in which a grammatical notion is marked on a head or on a dependent of that head, for example, marking grammatical functions through verb-agreement (head-marking) vs. case on nominals (dependent-marking). Nichols also made a number of proposals regarding the areal distribution of typologically variant properties. **Hogg (1992–2001)** is the invaluable *Cambridge History of the English Language*, a six-volume work which provides extremely detailed information about every aspect of the history of the language, from its Germanic and Indo-European origins to the present day. **Jasanoff (2004)** is a description of Gothic, a contribution to the *Cambridge Encyclopedia of Ancient Languages*, which gives descriptions of all languages known to have existed prior to 500AD for which a reasonable amount of data is available.

Phonological theory and phonological change

Chomsky and Halle (1968) is the classic exposition of generative phonology. It is notable for the system of distinctive features proposed, for the explication of the functioning of an ordered system of phonological rules, for the postulation of the levels of ‘systematic phonetics’ and ‘systematic phonemics’, for the analysis of the cyclic nature of stress-assignment in English, and for the markedness conventions connected to the evaluation metric discussed in §3.4. **Kiparsky (1973)** is a treatment of the nature of rule-ordering in the standard model of generative phonology as put forward by Chomsky and Halle (1968), in which the Elsewhere Condition is put forward as a condition determining one kind of rule-ordering. The origins of this condition in the works of Pāṇini’s are explicitly acknowledged. **Kenstowicz (1991)** is a standard, comprehensive introduction to (pre-optimality-theory) generative phonology. **Martinet (1955)** is a classic

structuralist account of phonetic and phonological change, in which the fundamental idea is that sound changes arise from the interaction of economy (ease of articulation) and expressiveness (the need to make distinctions). **Lass (1992)** is a further contribution to the *Cambridge History of the English Language*, in which the phonology and morphology of Middle and Early Modern English are described in detail.

This page intentionally left blank

4

The dynamics of syntactic change

Introduction	291	4.4. Reconstruction	357
4.1. Gradualness	293	4.5. Conclusion to Chapter 4	376
4.2. The spread of syntactic change	315	Further reading	377
4.3. Drift: the question of the direction of change	340		

Introduction

The previous chapters have outlined the general approach to syntactic change that the principles-and-parameters approach makes possible. As we have seen, syntactic change is seen as changes in values of parameters of UG taking place through the process of first-language acquisition. The changes are driven by the nature of the parameter-setting device, which, as it were, tries to set parameters on the basis of the PLD as efficiently as possible. The drive towards efficiency of parameter-setting leads to both the preference for simplicity of postulated derivations or representations and the tendency to generalize the input: both of these preferences underlie the markedness values associated either with parameters or parameter systems, as we saw in some detail in Chapter 3. We also suggested, following an old

idea, that efficiency and expressivity may to some degree pull in different directions, causing systems, over the long term at least, to show a general equilibrium as far as markedness is concerned.

The general view of change, acquisition, and parameters that we have outlined up to now is the one that we will maintain in the remainder of the book. The goal of this chapter is to see to what extent this view can shed any light on, or may be challenged by, certain fairly standard observations concerning syntactic change. For the purposes of this chapter, the two most important aspects of the approach to syntactic change that we have described are: (i) that it is catastrophic, in the sense that a given parameter changes its value suddenly and irrevocably, at a given historical moment; and, (ii) that it is ‘internal’ to the language acquirer, and in principle entirely independent of the social, cultural, or historical environment of that acquirer; all that counts is the acquirer’s linguistic environment, i.e. whatever aspect of the PLD that leads to the change in the parameter value.

Both of the aspects of parametric change just described are entailed by the approach described up to now, and yet both are to a certain degree inimical to what is often assumed, explicitly or implicitly, in much work in historical linguistics. Language change very often seems to be gradual rather than in any way sudden or catastrophic, and ‘external’ forces of various kinds – social, cultural, and historical – have often been thought to be central to understanding the nature of language change. (Both of these points are forcefully made in Weinreich, Labov, and Herzog (1968).) Two important facets of gradualness are, on the one hand, the concept of lexical diffusion (the idea that changes may gradually ‘diffuse’ through the lexicon item by item, rather than affecting the entire grammar at once) and, on the other, the concept of drift (the idea that languages may change over very long periods in certain preferred directions). Both lexical diffusion and drift contrast sharply with, and might at least seem to be in conflict with, the hypothesized sudden, discrete nature of parametric change; we will discuss to what extent there truly is a conflict here in the first three sections of this chapter. A further aspect of parametric change is the fact that it is irrevocable; this has been argued by Lightfoot (1979; 2002a, b) to render syntactic reconstruction impossible. Since reconstruction has proven to be a powerful and revealing tool in historical phonology and morphology, it is natural to ask whether it can also be used in syntax. In §4.4 I will review Lightfoot’s arguments that the nature of parametric change makes reconstruction impossible; at the same time, I will briefly review some recent work on

quantifying linguistic relatedness, and how this may be adapted in the context of parametric theory for the purposes of syntactic reconstruction.

One difficult question related to ‘external’ aspects of change has to do with the fact that ‘successful’ changes – those which have actually led to the elimination of an earlier system in favour of an innovative one in the attested record – are not confined to individuals, but concern aggregates of individuals, i.e. speech communities.¹ In order for our approach to account fully for change we need to say something about how changes may spread through a speech community; this will be addressed in §4.2.

More generally, at least the first two sections of this chapter are an attempt to deal, in the context of the theoretical assumptions regarding syntactic change that have been made in the preceding chapters, with what Weinreich, Labov, and Herzog (1968: 98–9) refer to as the paradoxical relationship between language structure and language history, and which they claim (120–2) to have been recognized by de Saussure (1959) when the distinction between synchrony and diachrony was first made. How can we reconcile the postulation of discrete, homogeneous algorithmic systems like generative grammars (or structuralist-style systems of elements in opposition) which are properties of individuals, with the apparently gradual change of a language at a given historical period in a given speech community? We have touched on this problem occasionally in previous chapters, but now it is time to try and deal with it systematically.

4.1. Gradualness

4.1.1. *Introduction*

As I mentioned in the Introduction, language change, including syntactic change, appears to be gradual. For example, Kroch (2000: 719) says: ‘[s]tudies of syntactic change which trace the temporal evolution of the forms in flux universally report that change is gradual’. He goes on to say:

¹ Strictly speaking, we should use a different term here, since ‘speech community’ could be taken to exclude Deaf communities which make use of sign language (one of which we will discuss in §5.4). However, I will continue to use this term owing to its familiarity, bearing in mind that it really refers to a linguistic community without prejudice as to the linguistic modality used in that community.

[b]efore the rise of generative grammar, this sort of gradualness was taken for granted. Syntactic change, once actuated, was conceived primarily as a slow drift in usage frequencies, which occasionally led to the loss of some linguistic forms. New forms, whether they entered the language as innovations or borrowings, would normally affect the language only marginally at the outset and then, if adopted by the speech community, would spread and rise in frequency.

Harris and Campbell (1995: 48) state that ‘syntactic change may be considered gradual in a number of respects’, although they also point out that ‘it is not useful to consider syntactic change in terms of a dichotomy between gradual vs. abrupt’ (*ibid.*). Moreover, we noted in §2.5.2 that Vennemann considers that certain changes may take millennia to come to completion, in commenting that ‘a language may become fairly consistent within a type in about 5,000 years (for example, English)’ (1974: 353).² Further, Haspelmath (1998: 344) refers to grammaticalization as ‘gradual unidirectional change ... turn[ing] ... lexical items into grammatical items’.

The idea that change is gradual arguably has two sources: on the one hand, the time course of change is impossible to pin down to a single historical moment. On the other hand, the emphasis on gradualness may stem in part from the influence that both geology – in the form of the uniformitarian thesis (introduced in §2.4), which was taken over from the work of the pioneering geologist Lyell (1830–4) – and Darwin’s theory of biological evolution have had on historical linguistics since the nineteenth century (see Lightfoot (1999: 41); MacMahon (1994: 334ff.) on the relation between Darwinian thinking and historical linguistics; Morpurgo-Davies (1998: 190–2, 196–7, 219, n. 2, 233, 282) on the influence of Darwin and Lyell on nineteenth-century linguists.)³ Both geology and the Darwinian approach to evolution emphasize the importance of gradual change over very long periods, although in the theory of evolution, ‘punctuated equilibrium’ has more recently become an influential point of view; see Eldredge and Gould (1972).

As Kroch points out, syntactic changes are not traceable to a specific historical moment, in the way that historical events such as the end of the First World War or the fall of the Byzantine Empire are. Instead, syntactic

² This comment really relates to the question of ‘typological drift’, which we will look at in §4.3. There is nonetheless a clearly implicit notion of gradual syntactic change.

³ Labov (1994: 23–4) insightfully points out that uniformitarianism need not imply gradualism in historical linguistics, although for a long time it was thought to do so in geology.

changes seem to be temporally diffuse. For this reason, it may be thought that the entailment of the approach to syntactic change that parametric change must be abrupt, sudden, or, to use Lightfoot's (1979, 1991) terminology, catastrophic, is problematic. However, further reflection reveals that there are compelling conceptual reasons to view syntactic change as sudden, in exactly the way we would expect given what we have seen up to now, and that there is in fact no strong empirical reason to consider that things should be otherwise, once a range of additional factors, both grammatical and sociolinguistic, are fully taken into consideration.

There are two main reasons why a gradualist view of syntactic change is conceptually untenable, given the general view of syntactic change that I have been propounding here. The first was lucidly articulated in Lightfoot (1979). Discussing the idea that syntactic change might be a form of drift (see §4.3), Lightfoot points out that no change involving more than a single generation is possible, given an acquisition-driven view of change:

Languages are learned and grammars constructed by the individuals of each generation. They do not have racial memories such that they know in some sense that their language has gradually been developing from, say, an SOV and towards an SVO type, and that it must continue along that path. After all, if there were a prescribed hierarchy of changes to be performed, how could a child, confronted with a language exactly half-way along this hierarchy, know whether the language was changing from type x to type y , or vice versa?

(Lightfoot 1979: 391)

(We have already touched upon this argument of Lightfoot's in relation to word-order change in §2.5.2, but the point applies to all types of syntactic change). As Lightfoot points out, this raises a seemingly insuperable problem for the postulation of any notion of typological drift, a matter we return to in §4.3.

The second point is a very simple, but conceptually powerful, one: parameter settings, like all other formal entities in generative grammar (and recall that parameter settings are no more than values of given formal features, or perhaps manifestations of such features – see (54) of Chapter 3) are discrete entities. As we pointed out in the initial discussion of parameters in §1.1, 'clines, continua, squishes and the like are ruled out'. Therefore gradual change from one value of a parameter to another is simply impossible: a parameter must be in one state or the other; it cannot be in between (cf. Kroch's (1989: 201) remark that 'the change from one grammar to another is necessarily instantaneous'). This is really a matter of

logic, not linguistics: the Law of the Excluded Middle ($p \vee \neg p$) is the relevant concept. Any system which appears to be ‘semi-null subject’ or ‘tendentially VO’ (and, of course, plenty do, as one simply surveys the data) must be analysed as being one thing or the other; strictly speaking no system can be in a state intermediate between two parameter values. This conclusion holds for any approach to linguistics which makes use of discrete categories, for example, grammatical categories such as verb, noun, etc., or phonemes of the usual type. Hockett (1958: 456–7), quoted in Weinreich, Labov, and Herzog (1968: 129) makes a similar observation when he points out that ‘[s]ound change itself is constant and slow. A phonemic restructuring, on the other hand, must in a sense be absolutely sudden’. No special claim is being made about parametric systems of the sort being assumed here in this respect. Hale (1998: 3) makes a related point in defining change as a set of differences between two grammars. He points out that this means that ‘“change” therefore has no temporal properties – it is a *set* of differences’ (emphasis in original). Since gradualness is usually intended as a temporal concept in discussions of language change, a consequence of Hale’s views is that change, as he defines it (and this definition applies to parametric change), cannot be gradual.

But it is undeniable that the ‘mirage’ (to use a term attributed to Mark Hale by Lightfoot (1999: 88)) of gradualness is very apparent to us in our observation of the phenomena. There are several reasons for this, many of which have been pointed out (see Harris and Campbell (1995: 49); Hopper and Traugott (2003: 229, 232); Roberts (1993a: 204); Willis (1998: 43–8)). The most obvious one concerns the time course of change: changes do not appear to be instantaneous in the historical record. Instead, they typically follow an S-shaped curve when frequency of occurrence of a new form vs. an old one is plotted against time. (This observation is due originally to Osgood and Sebeok (1954: 155); see Kroch (1989: 203); Weinreich, Labov, and Herzog (1968: 113, n. 20)). That is, changes start slowly, gather speed, and then taper off slowly again. Denison (1999) dubbed this pattern of change ‘slow, slow, quick, quick, slow’ (see also Denison (2003)). The S-curve is shown in Figure 4.1.

Kroch (1989) reconciled this fact about the time course of change with the discrete nature of grammars by postulating the ideas that grammars may ‘compete’, and that one replaces another at a constant rate. We will consider Kroch’s proposals in detail below and in the next section.

Alongside competing grammars, the other factors giving rise to the mirage of gradualness fall into two main kinds. First, there are sociolinguistic

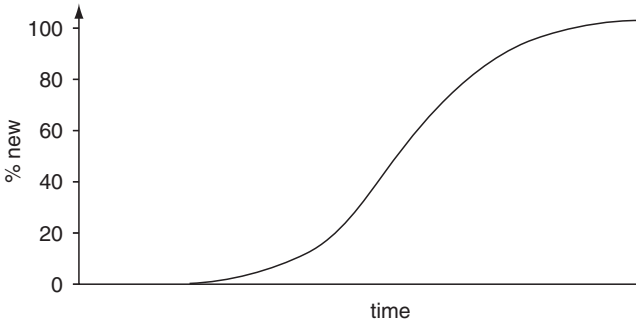


Figure 4.1 An idealized graphical change (from Denison (2003: 56))

factors: these include the often highly artificial and restricted register of surviving texts, variation in individual usage conditioned by style, accommodation, etc., and variation in a speaker's or writer's age, gender, class, etc. We will consider some of these issues in slightly more detail in the next section. Second, there are factors which arise from the nature of the grammatical system. These include lexical diffusion, extremely fine-grained parameters ('microparameters') whose precise nature may have gone unrecognized by analysts, and true optionality. All of these factors, both sociolinguistic and grammatical, are, as far as anyone knows, independent of one another. Their combined effect is the mirage of gradualness. In other words, these factors may cushion the effects of an instantaneous, discrete, structural change in the historical record.

Hale (1998: 5–7) distinguishes 'change' which, as we saw above, has no temporal properties, from 'diffusion', by which he means the process whereby a change spreads 'from the innovator to (a subset of) those with whom the innovator comes into contact' (Hale 1998: 5). As such, diffusion may have temporal properties, and the evidence of change is preserved in the historical record in the form of 'diffusion events', as we witness one system replacing another over time. Once again, no special claim is being made about parametric systems in this respect which would not be made about any formal system making use of discrete entities.

4.1.2. *Lexical diffusion*

Let us now look at the grammatical factors that may 'cushion' change in the sense described above in more detail, beginning with lexical diffusion. The

term lexical diffusion refers to the idea that a change may spread through the lexicon gradually, perhaps in the last analysis one word at a time. This view has been proposed, somewhat controversially, for sound changes (see in particular Labov (1994: 421ff.)). It is arguably more widely accepted as an operative concept in syntactic change, at least according to Harris and Campbell (1995: 107). We have seen an example of lexical diffusion in syntax in our discussion of changes to psych verbs in ME in §2.3.2. There, we saw how one type of OE psych construction – that in which the Experiencer is marked dative – was lost, with the original dative Experiencer being reanalysed as the subject and the original nominative Theme as the direct object. Thus, for example, OE *lician*, the ancestor of NE *like*, used to appear in examples like (1) (repeated from (36) of Chapter 2):

- (1) hu him se siga gelicade
 how him-Dat the victory-Nom pleased
 ‘how the victory had pleased him’
 (*Or* 84.32; Denison 1993: 72)

We observed that this verb has undergone a redistribution of its thematic roles (although not really a change in meaning, since the core meaning has involved causing pleasure all along), associated with the loss of the construction in (1). Allen (1995) argues convincingly that the changes in the psych verbs were changes in lexical entries of individual verbs, which diffused through the lexicon over a considerable period. She argues (221ff.) that the beginnings of this change may be discerned in the optional assignment of dative case to the Experiencer arguments of certain verbs in OE and that the change was completed only by approximately 1500. One piece of evidence for this is the innovation of new verbs with dative Experiencers as late as the fourteenth century; Allen (1995: 250ff.) gives examples of this with *ought* and *need*. So lexical entries of individual verbs may change in a piecemeal fashion: the specification of categorial and semantic selection properties can be altered, and this can have syntactic and semantic effects. Of course, such changes may not be completely piecemeal, or totally dissociated from changes in other parts of the grammar; we saw in our discussion of the psych verbs that Allen argues that the loss of morphological case distinctions in ME facilitated but did not cause this change. We also saw there that the individual changes may involve reanalysis, but not, in this instance, parametric change (see the discussion of (46) in §2.3.2). It is also likely that changes diffuse through (perhaps

rather small) subclasses of lexical items; see Kiparsky (2003: 315ff.) for a discussion of this in relation to phonological change. The important point for our present purposes, however, is that lexical diffusion can definitely give rise to gradual change: if we consider the change in psych verbs in English as a single change, then we have to say that it went on essentially throughout the entire ME period. However, strictly speaking, looking at the grammatical systems underlying what we refer to as ME, there was a series of discrete reanalyses across a subclass of lexical entries; there is no reason to treat these as gradual (and, indeed, reason not to treat them as gradual, since they affected discrete formal entities: the features of lexical entries).

The existence of lexical diffusion as a kind of change distinct from parametric change raises an interesting theoretical question. Most minimalist work follows the general assumption put forward in Chomsky (2001: 2), following an original proposal in Borer (1984), that ‘parametric variation is restricted to the lexicon’, involving the specification of formal features of functional heads. As we saw in detail in the last chapter, and illustrated in our formulation of parameters in (54) there, we follow this general approach here. But if parameters are specified in lexical entries, how are these lexical entries different from the lexical entries of verbs, which, as we have seen, are subject to lexical diffusion? It is in fact quite reasonable to say that there is no relevant difference; the only difference is the category that the lexical entries are associated with. In the diffusion example just discussed, we are dealing with the fairly rich lexical entries of a particular class of lexical verbs, complete with formal and semantic features of various kinds. In the case of parametric variation, on the other hand, we are dealing with the lexical entries of functional items, which may be restricted to formal features of various types. (See Roberts and Roussou (2003: 229ff.) for discussion of the idea that functional categories tend to lack phonological and semantic features.) Hence change in formal features of a functional category may have a greater effect on the nature of that category than change in formal features of a lexical category. It also has a far greater effect on the derivations and representations produced by the grammar: changing a feature of C, T, or *v* will affect any clause, but changing a feature of a psych verb, for example, will have a considerably smaller effect on the language.

If we maintain that there is no difference in principle between the lexical entries of lexical categories and those of functional categories beyond the

fact that the latter tend to be relatively impoverished, then parametric change is just defined as change in formal features of functional categories, and lexical change – therefore possibly lexical diffusion – as the same thing happening to the lexical entries of lexical categories. This opens up the possibility of ‘lexical diffusion’ through the functional system: a series of discrete changes to the formal features of a set of functional categories taking place over a long period and giving the impression of a single, large, gradual change. We have in fact already encountered this: we can interpret this as being what Vennemann was referring to in stating that change in head-complement order might take several millennia to complete. (See the quotation from Vennemann (1974: 353) at the beginning of this section.) We can see this as change, arguably in the EPP features of functional categories, diffusing through the system of functional categories and following the markedness convention we proposed in §3.5.2 (57).

4.1.3. *Microparametric change*

The notion of lexical diffusion of features of functional categories as a mechanism of syntactic change raises the question of ‘microparametric’ variation. Kayne (2000: 3–9) discusses this concept in some detail, pointing out that the study of large numbers of very closely related grammatical systems is a reliable way of ultimately isolating the ‘minimal units of syntactic variation’ (6). We are assuming here that these are the formal features of functional categories. The possibility of the lexical diffusion of such features just adumbrated creates the expectation that we may observe very fine-grained synchronic variation, due to variation in minimally different features of functional heads, and, apparently gradual diachronic change as exactly those features alter their values.

A case in point, which is of considerable general interest, concerns the phenomenon of ‘auxiliary selection’ in Central and Southern Italian dialects. The term ‘auxiliary selection’ refers to the choice of the equivalent of ‘have’ or ‘be’ as the auxiliary for compound tenses such as the perfect. In Standard Italian, as in German and Dutch, the choice of auxiliary depends on the argument structure of the main verb: if the verb is transitive or an unergative intransitive, the auxiliary is ‘have’; if the verb is an unaccusative intransitive, passive or marked with the clitic *si* (which has various uses we

will not go into here), the auxiliary is ‘be’. These facts are illustrated in (2) and (3):⁴

- (2) a. Maria ha visto Gianni. (transitive)
 Maria has seen-m.sg. Gianni
 ‘Mary has seen John.’
- b. Maria ha telefonato. (unergative intransitive)
 Maria has phoned-m.sg.
 ‘Mary has phoned.’
- (3) a. Maria è arrivata. (unaccusative intransitive)
 Maria is arrived-f.sg.
 ‘Mary has arrived.’
- b. Maria è stata accusata. (passive)
 Maria is been.-f.sg. accused-f.sg.
 ‘Mary has been accused.’
- c. Maria si è accusata. (reflexive)
 Maria SI is accused-f.sg.
 ‘Mary accused herself.’

(The distinction between unaccusative and unergative intransitives was presented in §2.3.1).

If we suppose that aspectual auxiliaries are always merged in *v*, then we can state the generalization for Standard Italian in terms of the notion of ‘defectivity’ of *v* in the sense of Chomsky (2001: 6–9), which we introduced in §2.3.1. Where *v* lacks what we called the ‘Burzio properties’ of Agreeing with the Case feature of the direct object and assigning a thematic role to the subject, it is defective. (In the earlier discussion we took it that *v* could also be absent, but we will see directly that this cannot be maintained here.) The generalization regarding Standard Italian auxiliary selection can be stated as in (4) (this is an updating and simplification of Burzio (1986: 53ff.)):

- (4) $v^*_{\text{Perfect}} = \text{have}; v_{\text{Perfect}} = \text{be}$

The ‘perfect’ feature here may well be a selectional feature, stating that the VP complement to *v* must be headed by a perfect participle; this feature

⁴ Where the auxiliary is ‘be’ in the examples in (3), the participle agrees with the subject. But where the auxiliary is ‘have’, the participle does not agree with the subject, and here shows the default masculine singular form. In the discussion to follow, I will largely leave aside the important question of the relation of participle agreement to auxiliary selection. This is purely for expository reasons: any full account of the facts must take this phenomenon into consideration.

must be present on all the instances of *v* discussed here, and so we will leave it out from now on.

As we said, *v** denotes a non-defective *v*, one capable of Agreeing with the direct object's Case and assigning an external thematic role to the subject. In other words, *v** is shorthand for two features: an external thematic role and a feature Agreeing with an internal Case. Moreover, if we assume that the direct object's structural Case feature Agrees with *v**'s person and number features (see §2.3.2), we see that defectivity is signalled by a rather complex set of features. Hence auxiliary selection is conditioned by this set of features.⁵

In languages like English or Spanish, where 'have' is the sole perfect auxiliary but where 'be' is used for the passive, the features of *v* which determine which auxiliary is merged (or which determine when 'have' is merged – see note 5), relate rather simply to voice and aspect. We can formulate the generalization as follows:

(5) *v*[active] = have.

Example (5) is clearly a simpler statement than (4), given that the notion of non-defectivity of *v* refers to a thematic feature and a set of person/number features. It is worth pointing out in this connection that both Spanish and English have lost an Italian-style auxiliary-selection system in their recorded histories (see Penny (1991: 142); Loporcaro (1998: 155ff.) on Spanish; Denison (1998: 135–8); and the references given there on English).⁶ We can see by comparing (4) and (5) that this change involved a simplification of the specification of the features of *v* required for the merger of 'have'.

⁵ 'Have'-selection is naturally seen as the marked option. There are several reasons to think this. First, 'have'-auxiliaries are cross-linguistically rather rare; in Indo-European they are not found in Celtic or Slavonic (with the exception of Macedonian (David Willis, p.c.)), or in Hindi (Mahajan 1994), for example. Second, any context where 'have' is found corresponds to one where 'be' can be found in some other language, but not vice versa. For example, 'have' is never, to my knowledge, the basic passive auxiliary. In all the languages mentioned here, this is always 'be'. Third, we can observe that the context for merger of 'have' has a longer description than that for 'be'. 'Be' can thus be considered to be the default auxiliary. Accordingly, all we need to do in order to give an account of auxiliary selection is specify the context where 'have' is merged. I will follow this practice for the remainder of this discussion.

⁶ Swedish (Faarlund 1994: 57), Portuguese, and Rumanian (Loporcaro 1998: 155) have also undergone this change.

Many dialects in Central and Southern Italy show a type of auxiliary-selection which is conditioned in quite a different way from that seen in Standard Italian. In these varieties, the argument structure of the verb is not the conditioning factor, but rather the person–number specification of the subject. There is much variation, but probably the most common pattern is that where ‘be’ appears with a 1st- or 2nd-person subject, and ‘have’ where the subject is 3rd person. This kind of system is found in Southern Lazio, Southern Marche, in Abruzzo, and in parts of Northern Apulia (Rohlf 1969: 123). There are a number of variations on this pattern, which I will not go into here (but cf. note 8 below).⁷ The microvariation found in these varieties is considerable. A particularly good example is the variation in the Neapolitan area discussed by Ledgeway (2000: 185ff.). According to Ledgeway, Literary Neapolitan has a system like that of Standard Italian, where auxiliary-selection is determined by argument structure. Urban Neapolitan, on the other hand, has a system like that of English or Spanish, where ‘have’ is always found in the active. Third, peripheral varieties (Torre del Greco, Torre Annunziata, Pompei, Sorrento; see Ledgeway (2000: 192) and the references given there) show the pattern where ‘be’ appears where the subject is in the 1st or 2nd person and ‘have’ where it is in the 3rd person. These systems are illustrated in (6)–(8) (Ledgeway’s (1)–(3), 186):⁸

⁷ On auxiliary selection in Central and Southern Italian dialects, see Cocchi (1995, Chapter 4); Kayne (2000: 115–7); Ledgeway (2000, Chapter 6); Lopocaro (1998); Manzini and Savoia (2005) Tuttle (1986); and the references given there.

⁸ Ledgeway discusses two further varieties: Procidano, where the choice of auxiliary is apparently determined by tense, and an obsolescent urban variety, where auxiliary selection is ‘determined by a combination of grammatical person and clitic-doubling’ (Ledgeway 2000: 186):

- (i) Procidano:
- a. Hó visto a Ciro/arrevèto.
I-have seen A Ciro/arrived
‘I have seen Ciro/arrived.’
 - b. Fove visto a Ciro/arrevèto.
I-had seen A Ciro/arrived
‘I had seen Ciro/arrived.’
- (ii) Obsolescent urban dialect:
- a. Aggiu visto a Ciro.
I-have seen A Ciro/arrived
‘I have seen Ciro/arrived.’

- (6) Literary Neapolitan:
- a. Aggiu visto a *Ciro*. (transitive)
 I-have seen A *Ciro*
 ‘I have seen *Ciro*.’
- b. *So*’ arrevato. (unaccusative)
 I-am arrived
 ‘I have arrived.’
- (7) Urban Spoken Neapolitan:
- a. Aggiu visto a *Ciro*. (transitive)
 I-have seen A *Ciro*
 ‘I have seen *Ciro*.’
- b. Aggiu arrevato. (unaccusative)
 I-have arrived
- (8) Peripheral varieties:
- a. *So*’ visto a *Ciro*/arrevato. (1st/2nd person subject)
 I-am seen A *Ciro*/arrived
 ‘I have seen *Ciro*/arrived.’
- b. *Ha* visto a *Ciro*/arrevato. (3rd-person subject)
 s/he-has seen A *Ciro*/arrived
 ‘s/he has seen *Ciro*/arrived.’

We can follow Ledgeway in taking the urban variety to be innovative in relation to the literary one. Here we see the same change taking place as that which has happened in English, Spanish, and the other languages mentioned in note 5. As we said above, this can be seen as a simplification in the environment for merger of ‘have’. The peripheral varieties also represent an innovation (see Tuttle (1986); Bentley and Eythórssen (2003); and Rohlf’s (1969: 123) comment that dialects with this system

- b. ‘*O so*’ visto a *Ciro*.
 him I-am seen A *Ciro*
 ‘I have seen *Ciro*.’
- c. (*L*)*ha* visto a *Ciro*.
 him he-has seen A *Ciro*
 ‘He has seen *Ciro*.’

In (iib, c) we see clitic-doubling of the direct object. The *a* marker preceding the direct object here and in (6)–(8) is comparable to that found in Spanish, and is another widespread feature of Central and Southern Italian dialects – see Ledgeway (2000, Chapter 2) and the references given there. The contrasts here show that the person-driven auxiliary alternation only appears where the direct object is clitic-doubled.

show ‘a very strange extension of *be* in place of *have* with transitive verbs’ (‘una stranissima estensione di *essere* in luogo di *avere* coi verbi transitivi’ – my translation) <Tuttle, Vincent>. We can define the context for merger of ‘have’ in *v* as in (9):

(9) $v[\text{active}, 3\text{pers}] = \text{have}$.

We can again observe that (9) is simpler than (4), although more complex than (5).⁹ Hence in these dialects (4) has developed into (9).¹⁰

We see then that the variation in auxiliary-selection in the Central and Southern Italian dialects can be reduced to variation in the features of *v* associated with merger of the ‘have’-auxiliary; this is a case of complex microparametric variation which can be handled in terms of the general approach to parametric variation assumed here. Alongside synchronic microvariation in auxiliary-selection (witness Ledgeway’s description of the variation in the Neapolitan area), we find seemingly gradual diachronic change: this is a further instance of apparent gradualness which can be reduced to discrete parameter settings, i.e. the formal features of functional categories.

4.1.4. Formal optionality

The final way in which apparent gradualness may come about as a consequence of properties of the grammatical system itself has to do with optionality. If grammars allow for true formal optionality, with no semantic consequences, then surface variants may exist (without grammar competition, since by assumption there is just one grammar underlying the variants). Random fluctuations in usage may then give rise to an impression of gradual change. But in fact in such a case there would be no

⁹ In the obsolescent urban dialects illustrated in (ii) of note 8, we have the system in (5) where there is no object-clitic doubling, and that in (9) where there is an object clitic. This implies the disjunction of (9), with the relevant contextual specification, and (5). It is therefore unsurprising that this system developed into that relying solely on (5).

¹⁰ There is an important complication associated with (9). The [3pers] feature Agrees with the subject. Being associated with *v*, we would expect it to Agree with the object. Here, we must invoke the analysis of ergativity put forward in Müller (2004b). See Box 4.1.

BOX 4.1 Ergative case marking

Ergative case/agreement systems differ from nominative-accusative ones in that in such systems the subject of a transitive verb receives a special case or agreement marking, while the subject of an intransitive has the same case/agreement marking as the direct object of a transitive. This pattern is illustrated by the following examples from Basque, taken from Comrie, Matthews, and Polinsky (2003: 45):

- (1) a. Ni-k neska ikusten dut.
 I-ERG girl see AUX
 ‘I see the girl.’
 b. Ni etorri naiz.
 I come AUX
 ‘I have come.’

The subject of the transitive verb in (1a) is marked with the ergative case, while the object has no overt case marking, just like the subject of the intransitive in (1b). (This form is often referred to as the absolutive.) This pattern contrasts with the familiar nominative-accusative one where subjects of transitives and intransitives pattern alike concerning case and agreement, while objects of transitives show different marking. The diagram in (2) illustrates the difference between the two types of system in schematic form:

(2)

	<i>Subject of transitive</i>	<i>Object of transitive</i>	<i>Subject of intransitive</i>
NOM-ACC system	NOM	ACC	NOM
ERG-ABS system	ERG	ABS	ABS

Müller (2004b) argues that the contrast between ergative case marking/agreement patterns and nominative-accusative patterns derives from a choice in the order of operations in a transitive clause when the derivation reaches *v*. Here, *v* may either Agree with the direct object, or the subject may be merged. If Agree precedes Merge, *v*'s features Agree with the Case feature of the object, and the subject, once merged, must Agree with T. This gives rise to a nominative-accusative system. On the other hand, if the subject is merged first, it can Agree with *v* and the direct

object will then Agree with T. (This conclusion requires one or two technical adjustments to the definition of Agree compared to that given in §1.4.1, which I will leave aside here.) This gives rise to an ergative system. It seems that we must assume that v's [3Pers] feature Agrees with the subject in the Central and Southern Italian dialects discussed in the text, as in an ergative system. On further similarities between these dialects and ergative systems, see Mahajan (1994); Manzini and Savoia (2005). The object's Case feature, however, may be checked by participle agreement (which is widely found in Central-Southern Italian dialects (Manzini and Savoia 2005, II: 745).). It is possible that T's features Agree only with those of the auxiliary; this is actually one possible technical instantiation of the null-subject parameter. This may explain why no Germanic variety is attested with the system in (9) (although in Germanic we find the systems in (4) and (5), and change from one to the other), since no Germanic language is a null-subject language. This might be a further, potentially very interesting, case of parameter interaction of the general kind characteristic of networks of parameters, as discussed in §3.5.3. D'Alessandro and Roberts (2006) develop this idea further in relation to the Eastern Abruzzese dialect.

grammatical change at all, simply a difference in the use of one or another form that grammar makes available. I will return to this last point in the next section, looking at how social factors may interact with formal, grammatical options.

Optionality has been seen as problematic in the context of minimalist syntax, since a central tenet in minimalism has been the idea that formal operations only apply when forced to: they are either obligatory, for example in the presence of the relevant triggering feature (such as Move being triggered by an EPP feature, or Agree by the presence of uninterpretable features), or, where the triggering feature is absent, they are impossible. This kind of approach appears to leave little room for optionality, beyond, of course, optionality in the choice of lexical items.¹¹

¹¹ Of course, one could always postulate optional EPP or uninterpretable features. But if these are the features which define parameters, and if different grammars are defined in terms of whether and how they differ in parameter values, then postulating optional features in this way is equivalent to postulating different grammars. Our concern here, however, is with the possibility of optionality in a single grammar.

However, it is not quite correct to conclude that true optionality within a single grammar is absolutely impossible. Recent work by Biberauer (2003); Biberauer and Richards (2006); Biberauer and Roberts (2005a); Richards and Biberauer (2005) has shown that in the context of the recent versions of minimalism, i.e. those presented in Chomsky (2000; 2001; 2004; 2005a, b), true optionality is a technical possibility. Moreover, Biberauer and Richards (2006) demonstrate that this possibility is in fact attested.

Briefly, the technical possibility of optionality derives from the dissociation of the operation eliminating uninterpretable features, Agree, from the triggering of movement by means of EPP features. (In earlier versions of minimalism, for example, Chomsky (1993; 1995), these were conflated.) One of the principal ways in which movement is triggered is in the case where the Probe of Agree has an EPP feature. In that case, the Goal of Agree is attracted to the Probe. However, and this is the central point in the present context, nothing prevents the category which moves from being larger than the Goal; all that the system requires is that the Goal either be exactly what is moved, or be contained in what is moved. More generally, movement involves the abstract configuration in (10):

(10) ... X[+EPP]_{PROBE} ... [YP ... Z_{GOAL} ...] ...

In this situation, in principle it does not matter whether Z or YP moves to X. Therefore the possibility of true optionality arises.¹² This is purely formal optionality, with no semantic consequences at all.

Biberauer and Richards (2006) discuss a number of cases where exactly this kind of optionality appears to obtain. One such case concerns wh-movement in Russian. In this language, a wh-pronoun may move alone (off the ‘left branch’ of the wh-DP) or the entire wh-DP may move:

- (11) a. Č’ju knigu ty čital?
 whose book you read
 b. Č’ju ty čital knigu?
 whose you read book
 ‘Whose book did you read?’
 (Biberauer and Richards 2006, (28))

¹² It may also be the case one or other option is excluded as a parametric choice; see Chapter 3, note 20, but this choice has to be determined by something else in the grammar or we would again have two grammars – see the previous note. Richards and Biberauer (2005) and Biberauer and Richards (2006) discuss this point in detail, and show how optional pied-piping in the Germanic languages which have it is determined by independent aspects of the grammar.

Here C triggers movement of a wh-phrase to its specifier; in (11a) the wh-element pied-pipes the larger DP, *č'ju knigu* ('which book'), while in (11b) just the wh-expression *č'ju* moves. In English (and many other languages), the equivalent of (11b) is ungrammatical (**Whose did you read book?*). Biberauer and Richards propose that this is because English wh-expressions are determiners, i.e. Ds, while wh-movement must involve movement of a phrasal category. In Russian, wh-expressions like *č'ju* are Quantifier Phrases (QPs), and so are able to move alone to SpecCP.

So we see that grammars may, under certain conditions, allow true optionality. In that case, if the two variants are attested at gradually varying frequencies over time, no grammatical change is at work, but simply a gradual shift in choice of options. A potential example of the diachronic situation, unfortunately lacking in full documentation for the relevant period, comes from the history of Greek. Biberauer and Richards (2006: 22) show that Ancient Greek displayed the same options as those seen in Russian regarding wh-movement. However, Modern Greek is like English in disallowing the equivalent of (11b). Their data and the account of the development in Greek is based on Mathieu and Sitaridou (2005), summarized in §1.5.2; see (124), (125) there. They argue that this indicates that the Greek wh-expressions have been historically reanalysed from QP to D, in line with the proposals in Mathieu and Sitaridou (2005) and Roberts and Roussou (2003: 161–7). If, at some intermediate stage of Greek, a gradual preference for the equivalent of (11a) over (11b) is attested, this would be exactly the case in point. Horrocks (1997: 224) dates the reorganization of the article system to spoken Byzantine Greek, and so in principle we might expect to find this gradual shift towards a preference for (11a) in this period, or just before (see also Manolissou (2001)). It is important to see that while the grammar, to use Biberauer and Richards' formulation, 'doesn't mind' which option is taken, other factors – including those of a sociolinguistic nature – may determine the choice made by a given speaker at a given historical moment. I will come back to this last point in the next section.

4.1.5. *The Constant Rate Effect*

Finally, let us return to Kroch's seminal work on the Constant Rate Effect and consider the proposal in more detail. Kroch (1989) identified the Constant Rate Effect for the first time. The Constant Rate Effect claims

that ‘when one grammatical option replaces another with which it is in competition across a set of linguistic contexts, the rate of replacement, properly measured, is the same in all of them’ (Kroch 1989: 200). Kroch goes on to point out that ‘the grammatical analysis that defines the contexts of a change is quite abstract’ (201), and that we must therefore ‘look for causes of change at more abstract levels of structure’ than simply contextual effects (239).

Kroch illustrates the Constant Rate Effect with a number of cases, the most striking of which is the development of periphrastic *do* in Early Modern English. Since this is connected to the change in the value of the V-to-T parameter discussed in §1.3.2, §2.1.5, and §3.3.2, let us now look in detail at this.

The starting point of Kroch’s demonstration of the Constant Rate Effect is the observation we reported above that changes tend to follow an S-shaped curve when the frequency of new vs. old forms is plotted against time (see Figure 4.1 above). The S-curve can be mathematically modelled by a function called the logistic. This function permits one to determine, as a linear function of time, the values of two quantities s and k . The first of these, s , is the ‘slope’ of the function and ‘hence represents the rate of replacement of the new form by the old’ (Kroch 1989:204). The second, k , is the ‘intercept’ value, which measures the frequency of the new form at some fixed point in time. What is of interest for the Constant Rate Effect are the values of s ; as Kroch says ‘[b]ecause fitting empirical data to the logistic function will allow us to estimate the slope parameters for each context of a changing form, we can determine, where sufficient data are available, whether the rates of change in different contexts are the same or different’ (Kroch 1989: 205–6). The Constant Rate Effect summarizes the empirical result that they are the same.

Kroch bases his study of the rise of periphrastic *do* on Ellegård’s (1953) survey of the instances of this element in texts for the period 1425–1600. Ellegård’s data is summarised in the graph in Figure 4.2.

As Kroch (1989: 223) points out, the curves are approximately S-shaped up to Ellegård’s Period 7, 1550–75. Kroch (1989: 224) observes that it ‘seems plausible to hypothesize that the point of inflection in Period 7 corresponds to a major reanalysis of the English auxiliary system’; this is the loss of V-to-T movement. Kroch applies the logistic to the data underlying Ellegård’s graph in Figure 4.2 and arrives at the following values for the slope parameter for Periods 1–7 (the contexts are those employed by Ellegård):

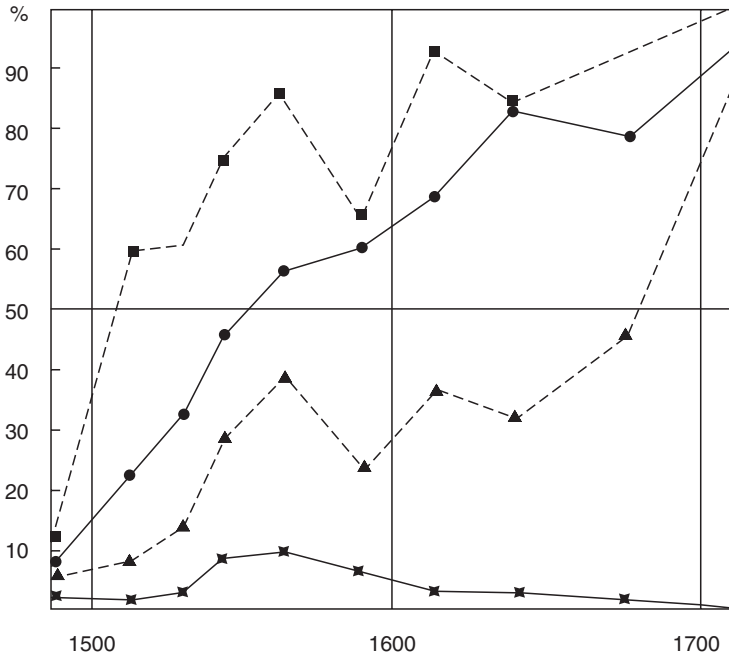


Figure 4.2 Auxiliary *do*. Percentage of *do* forms in different types of sentence, 1500–1700.

Upper broken line: negative questions.

Upper solid line: affirmative questions.

Lower broken line: negative declarative sentences.

Lower solid line: affirmative declarative sentences.

Adapted from Ellegård *The Auxiliary Do*. University of Gothenburg, 1953 (see also Lench (1989: 223)).

(12) Negative declaratives: 3.74

Negative questions: 3.45

Affirmative transitive adverbial and yes/no questions: 3.62

Affirmative intransitive adverbial and yes/no questions: 3.77

Affirmative wh-object questions: 4.01

Kroch points out that the single best slope for all five contexts has a value of 3.70 and that the probability of random fluctuations giving rise to the deviations from this value observed in (12) is .95 ($\chi^2 = .504$). As he says, the results ‘support the hypothesis that the slopes of the curves are underlyingly the same and the observed differences among them are random fluctuations’ (225). This shows that the grammar with *do*-insertion is replacing the grammar with V-to-T movement at a constant rate in all

contexts, a conclusion which justifies the postulation of something like the V-to-T parameter of §1.3.2, as well as the idea that this parameter was changing its value in the sixteenth-century. Of course, we still need to understand what it means for the grammars to compete; we will come back to this in the next section. Kroch goes on to confirm this conclusion by comparing the rate of replacement of V-Adv order with Adv-V order (the adverb in question is *never*); recall that this is a further indication of the loss of V-to-T movement.¹³ It emerges that the slope of the curve plotting the replacement of V-Adv by Adv-V has very nearly the same value as that plotting the rise of *do* in questions and negatives. He concludes ‘we have here substantial evidence that all contexts reflecting the loss of V-to-I raising [i.e. V-to-T raising – IGR] change at the same rate’ (229).

A well-known feature of sixteenth-century English, which we commented on in §2.1.5, is the fact that periphrastic *do* could also appear in positive declarative contexts, where it is no longer available (without emphasis) in Modern Standard English. Kroch shows that this occurrence of *do* changed at the same rate as *do* in the other contexts up to Period 7. After this, however, the use of positive declarative *do* declines, while its use increases in the other contexts.¹⁴ Kroch suggests that this is because positive declarative *do* starts to compete with ‘affix hopping’ (i.e. the Agree_φ relation between T and V – see §1.4.1) once V-to-T is lost, i.e. after Period 7, and ‘loses’ this competition. Hence the development of positive declarative *do* does not follow an S-curve and does not, after Period 7, change at a constant rate with the other contexts for periphrastic *do*.

Although not required by Kroch’s assumptions, we can take ‘grammatical option’ in the statement of the Constant Rate Effect to refer to a parametric option. This is justified by the fact that the Constant Rate Effect can reveal clustering of surface changes of the type that clearly indicate change in a single underlying parameter. As we have just seen, in Kroch’s discussion of the rise of periphrastic *do*, we are witnessing the

¹³ There is a complication in that ME allowed Adv-V order as well. See Kroch (1989: 226–7) for a discussion of the problem and a proposed solution.

¹⁴ Although not monotonically: the interrogative contexts continue to change together and to increase in frequency, but negative declarative *do* decreases for the period immediately after 1575, and only picks up again after 1650. Kroch suggests, very plausibly, that this was caused by independent changes in the syntax of *not*. This is discussed further in §4.3 below and in Roberts (1993a: 303–5).

change in the value of the V-to-T parameter. We can thus formulate the Constant Rate Effect in parametric terms as follows:

(13) Value v_j of parameter P_i replaces value $v_{i \neq j}$ at a constant rate.

It is important to see that the Constant Rate Effect implies that one parameter setting will be in competition with another for a certain period; I will come back to this notion of competing grammars (differing in the value of at least one parameter) in more detail in the next section.

The Constant Rate Effect is of great interest for two reasons. First, it reduces the observed gradual replacement of one form by another to competing grammars. Since distinct grammars are distinct entities, defined as differing from one another in at least one parameter value, the appearance of gradual change really reflects one discrete entity gradually replacing another, either in a speech community or in an individual. The gradualness is then an effect either of spread through a population, or of some factor causing an individual endowed with competence in more than one grammar to access one of these grammars more readily than the other(s) over time. In these terms, (13) can be reformulated as follows:

(13') In a given speech community/individual, grammar G_i with parameter P_i set to value v_i replaces grammar G_j with parameter P_i set to value $v_{j \neq i}$ at a constant rate.

One might wonder why (13') should hold. It is unlikely to be a fact about the grammars themselves. Instead, it is plausible that it may be a fact either about speech communities or about the ways in which individuals choose among grammars available to them. As such, it may be attributable to sociolinguistic factors or to the dynamics of populations, or both factors acting in tandem.

The second reason that the Constant Rate Effect is of great interest is that it can provide direct evidence of the clustering effect we expect from parametric change. The 'different contexts' Kroch refers to are the different surface manifestations of a given parameter setting. We saw this in detail with the example of periphrastic *do* above. In principle, it could be repeated with other examples of parameters. (Kroch (1989: 210–5) in fact discusses the loss of V2 in French; see §1.3.2.) Willis (1998: 47) makes this point regarding the Constant Rate Effect, and observes that what he calls the 'uniform diffusion' of the consequences of a parametric change manifest themselves as the Constant Rate Effect. It is worth pointing out that a statistical treatment of the data connected to a parametric change of at

least the level of sophistication and detail of Kroch's study may be needed in order to demonstrate the uniform diffusion of a parametric change; the textual record may very well not wear the effects of parametric change on its sleeve.

4.1.6. *Conclusion*

In this section we have seen several reasons to think that the observed gradualness in the time course of syntactic change is illusory. This is so since grammars rely on discrete entities, the relevant one for our purposes being parameter values, and because language acquirers can have no information as to 'ongoing changes' in the PLD to which they are exposed. The illusion of gradual change can be traced to a variety of factors, both grammatical and sociolinguistic. Here I have concentrated on the grammatical factors: lexical diffusion, microparametric variation, and true optionality. I also illustrated, following Kroch (1989), the Constant Rate Effect, showing how this reduces gradualness in syntactic change to extragrammatical grammar competition. It is now time to review the nature of this grammar competition in more detail, and to consider the sociolinguistic factors which influence the time course of syntactic change. More generally, we must address the role of 'external' forces (i.e. forces external to the language acquirer other than the PLD itself) in syntactic change.¹⁵

¹⁵ Maria-Teresa Guasti (p.c.) points out that there is evidence for gradual change, even in the usage of a single child, in first-language acquisition: she presents data showing the gradual time course of the loss of both early null subjects and root infinitives (see Figures 5.2, 5.3, and 5.4 in Guasti (2002: 165, 176)). It would be very interesting to know if the Constant Rate Effect can be observed in these cases. It may be possible to appeal to competing grammars in language acquisition, as suggested in Clark and Roberts (1993) and Yang (2000; 2002), in order to account for this. It is also possible that these facts further demonstrate that language acquisition is fundamentally different from language change, owing perhaps to the effects of maturation. If maturation can be shown to be at work in first-language acquisition, then the uniformitarian thesis does not hold in this domain, whereas we are following the standard assumption that it does in language change. (See the discussion at the end of §2.4.)

4.2. The spread of syntactic change

4.2.1. Introduction

The central question that I want to address in this section is: how does a parameter change spread through a speech community? An account of this is necessary in order to give a full description of the nature of ‘successful’ syntactic changes, i.e. those whose result is the complete replacement of an old grammar by a new one. The internalist account of parameter setting as being a consequence of aspects of the first-language acquisition process, which we described in the preceding chapter, gives us a very interesting perspective on change at the individual level, but, on its own, it cannot tell us about how changes affect speech communities. This point is really at the heart of what Weinreich, Labov, and Herzog (henceforth WLH) (1968: 98ff.) refer to as the paradoxes of historical linguistics; discussing Paul’s (1920) account of language change, they state that ‘[i]n isolating the language of the individual from the language custom of the group, Paul developed a dichotomy which was adopted by generations of succeeding linguists and which lies ... at the bottom of the twentieth-century paradoxes concerning language change’ (104).

We can put this issue in terms of the distinction between the actuation and the various aspects of the implementation of change as described by WLH (101–2). The most important aspect of implementation for our purposes is the transition problem: ‘the intervening stages which can be observed, or which must be posited, between any two forms of a language defined for a language community at different times’ (WLH: 101). The actuation of a change is the introduction of a novel form; the transition of a change is the spread of that form through a speech community. The approach to syntactic change as driven by first-language acquisition that we have advocated here arguably provides an account of the actuation of syntactic change (although WLH (145–6) make some critical remarks about Halle’s (1962) similar proposals regarding phonological change; I will return to these below), but says nothing about the transition problem. That is what I want to address here.

We begin by taking up again Kroch’s notion of grammars in competition. We will scrutinize this more closely, looking at grammar competition both in the the speech community and in the individual. In this context,

I take up the question posed at the end of the previous section: what are the Constant Rate Effect and S-curve describing change really properties of? Here I also introduce the concept of ‘syntactic **diglossia**’: the idea that individuals, and therefore speech communities, may synchronically instantiate several grammatical systems, with only minimal phonological and lexical variation. This leads naturally to a discussion of **code-switching**.

Diglossia – syntactic or otherwise – and code-switching are cases of what WLH (151) call ‘orderly differentiation’, which they take to be a key notion in understanding language change. Orderly differentiation can include both coexistent systems – languages or dialects in contact – or variation within a single system, and it may be keyed either to other aspects of the system or to sociological factors such as age, sex, class, and ethnicity.

4.2.2. *Orderly differentiation and social stratification*

A well-known example of orderly differentiation involving the social factors related to a sound change comes from the incidence of final and preconsonantal (r)¹⁶ in New York City (see Labov (1966); WLH (177–81)). Figure 4.3 shows the frequency of final and preconsonantal (r) (i.e. the pronunciation of *car* as /kar/ as opposed to /ka:/) in the casual speech of adult native New Yorkers in the mid-1960s. The horizontal axis plots age and socio-economic class; the vertical axis the percentage of (r).

As WLH (1968: 179) point out, this figure ‘shows an increase in the *stratification* of (r): the distance between the upper middle class and the rest of the population is increasing’. They further point out that the pronunciation of [r] ‘has evidently acquired the social significance of a prestige pronunciation’ for younger speakers, while for older speakers ‘there is no particular pattern to the distribution of (r)’ (178). Here we see an example of a linguistic variable taking on social value; Labov also demonstrated how speakers may be conscious of such social value.¹⁷

¹⁶ Sociolinguistic variables having to do with phonetic or phonological variation are commonly written in normal parentheses, as opposed to the square brackets used for phonetic transcription and the obliques used for phonemic transcription.

¹⁷ Labov’s original data, published in 1966, was collected in 1962 and was replicated by Fowler (1986). She found that, overall, little had changed in the social stratification of this variable in the intervening years. Labov (1994: 94) concludes that this change is still in the initial ‘slow’ change of the lower part of the S-curve, saying ‘at some point in the process we must expect a sudden acceleration’ (*ibid.*).

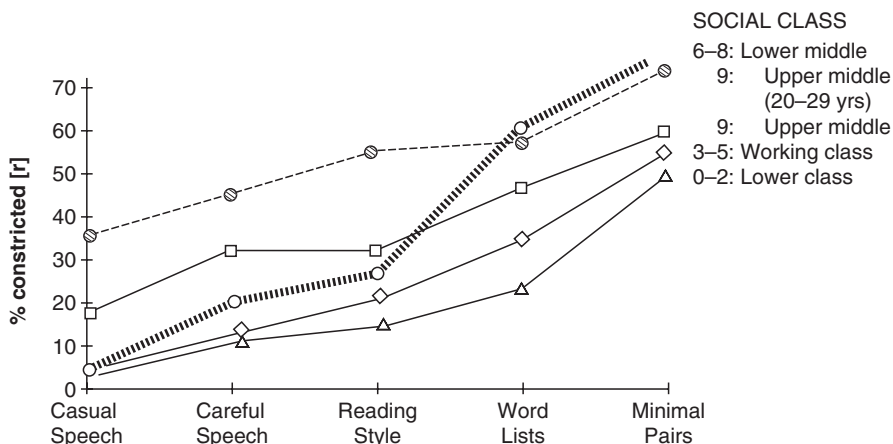


Figure 4.3 Social stratification of (r) in New York City (from Labov (1994: 87))

Labov (1966; 1972; 1994; 2001) has shown very convincingly how social factors influence the spread of sound change through a speech community, and so it is natural to ask whether they also influence syntactic change. In fact, one of Labov's pioneering studies concerned negative concord in African-American Vernacular English (AAVE). Labov (1972) makes a number of striking observations concerning negative concord, and other aspects of the syntax of negation, in AAVE. The most important observations are (i) that negative concord can manifest itself across clauses, in that formal, clausal negation in a complement or adjunct clause can be construed as an instance of negative concord with clausal negation in a higher clause, and (ii) negative concord can affect the subject. (Thus AAVE is a 'strict negative-concord' language in the terminology of Box 1.4 of Chapter 1.) These two phenomena are illustrated in (14):

- (14) a. Well, wasn't much I couldn't do.
(*Derek*; Labov 1972: 151)
- b. Down here nobody don't know about no club.
(*William T.*, 25, *Florida*; Labov 1972: 149)

The two phenomena combine to give rise to examples of the following type:

- (15) a. It ain't no cat can't get in no coop.
b. When it rained, nobody don't know it didn't.

Each of these examples actually contains a single logical negation ('No cat can get into a coop'; '... nobody knew it did'), the further expressions of formal negation being reflexes of the negative-concord system of this

variety. In the terms we introduced in §1.4, it must be the case that AAVE allows non-inverse negative Agree relations, and that clausal negation (usually expressed by an auxiliary combined with *n't*) is not obligatorily associated with an interpretable negative feature.

Labov points out that in examples like those in (15) we observe two changes. The first is that negative concord has 'lost its emphatic character' in AAVE, and the second is its spread into the new environment, in that it extends to clausal negation in lower clauses. The 'loss of emphatic character' is seen partly through the existence of other devices to indicate emphatic negation in AAVE, as follows:

- (16) a. Introduction of 'extra' quantifiers:
 She ain't in no seventh grade.
 b. 'Free floating negatives':
 But not my physical structure can't walk through that wall.
 c. Negative inversion:
 Ain't nobody on the block go to school.
 d. Concord with new quantifiers:
 Don't so many people do it.

Labov argues that even Standard English has negative concord in emphatic contexts (*pace* the rather simplified discussion in §1.4.1), and so the loss of the emphatic interpretation is an important change. It may be that this, along with the introduction of 'extra' negative elements, plays a role in the general development from Stage I to Stage II of Jespersen's Cycle, as discussed in §2.2. To speakers unfamiliar with AAVE, it is the extension of negative concord to clausal negation in lower clauses that is the really striking innovation.

Labov compares AAVE with two non-standard varieties of English spoken by whites: one in New York City and the other in Atlanta. Both have negative concord, but the New York variety does not allow negative concord on a lower verb; the Atlanta variety does but does not allow 'long-distance' negative concord otherwise. (These facts are summed up in Labov's Table 4.6, 193.) Moreover, negative concord is variable in these varieties, but not in AAVE, where non-concord is only found in contexts which can be construed as involving code-switching with the standard variety (Labov 1972: 184–5). All of this clearly shows that AAVE is a distinct grammatical system from Standard English (and from the non-standard white varieties), and that it may have undergone, or be undergoing, certain parametric changes which do not affect other varieties. And of

course this distinct variety has social value, since it is associated with a particular ethnic group. Syntactic variation and change can also, therefore, be associated with social value and social stratification.

4.2.3. *Grammars in competition*

Let us return now the question of grammars in competition, another potential case of orderly differentiation in the sense of WLH. In addition to the cases discussed in Kroch (1989), Santorini (1989; 1992; 1993) argues for grammar competition in word-order change in the history of Yiddish; Pintzuk (1991) very influentially argued the same for word-order change in the history of English (as we mentioned in §1.6.2); similarly A. Taylor (1994) reaches the same conclusion regarding word-order change between Homeric and Classical Greek. Finally, Fontana (1993) argues that the loss of V2 in Spanish involved grammar competition. These and a number of other studies have contributed a substantial body of work on syntactic change making use of the idea of competing grammars (see Kroch (2000: 720); Pintzuk (2003: 518)).

As we saw at the end of §4.1, the Constant Rate Effect can be seen as the consequence of the presence of two grammars in an individual or a speech community where one grammar is in the process of replacing the other – see (13'). In fact, the Constant Rate Effect is an aspect of the transition of a change: the logistic function tracks the fraction of the advancing form in relation to the slope and the intercept, and in relation to time. As Kroch (1989: 205) points out, the fraction of the advancing form 'jumps from zero to some small positive value in a temporal discontinuity'; this is the point of actuation. The actuation can be seen as the abductive reanalyses associated with parametric change, as described in the preceding chapters. The transition is the spread of the new parameter setting, with the associated reanalysed structures, through a speech community. So (13') is a generalization about the implementation of a change.

If the Constant Rate Effect, in the guise of (13'), is a generalization about the transition of a change, what is the S-curve? We can phrase this question in another way: following Kroch, in order to countenance the Constant Rate Effect at all, we must assume that more than one grammar may be present in a speech community at a given historical moment. If we are to abandon the idealization of an ideal speaker/hearer in a homogeneous

speech community (see Chomsky (1965: 3)), which has proven so useful for the purposes of linguistic theory, and instead consider ‘actual people in specific circumstances’ (Kroch 1989: 202), such as speakers of sixteenth-century English or contemporary Neapolitans controlling a range of dialects in addition to Standard Italian, then we simply must admit that a single speech community may feature more than one grammar. It seems, then, that we have no choice other than to adopt the idea of grammars in competition. However, we are really only forced to adopt the idea of *coexisting* grammars, and there is no reason – as far as the grammars themselves are concerned – why this coexistence should not be peaceful. To deny this would be to deny that an individual can be truly, natively bilingual, and this seems contrary to fact. So, why do we find change: the case where one of the grammars ousts the other one, in conformity with the S-curve? In other words, although the Constant Rate Effect entails competing or coexisting grammars, competing/coexisting grammars do not entail the Constant Rate Effect. As it stands, the competition/coexistence model explains neither the inception of the change nor its completion. We can consider the inception of the change to be actuation, and perhaps explain it along the lines described in the previous chapter, but this still leaves open the question of the completion of the change, i.e. the simple fact that one grammar eventually ousts the other. So the competition/coexistence idea needs supplementing if we are to arrive at an account of the transition of change.

One possibility would be to construe the actuation/implementation distinction as illusory, and to consider transition to be simply multiple cases of actuation (i.e. reanalyses and associated parameter change) across a population. The initial actuation could take place in a very small group of acquirers, for the kinds of reasons discussed in the previous chapter, and spread through the population as more and more acquirers are exposed to the innovative system. Assuming the innovative system to be favoured by the parameter-setting device, then we can understand why this system eventually ‘wins’. The S-curve is then seen as a property of population dynamics; cf. Kroch’s (1989: 203) observation that ‘[i]n the domain of population biology, it is demonstrable that the logistic governs the replacement of organisms and of genetic alleles that differ in Darwinian fitness’. (Recall that the logistic underlies the S-curve.) If we construe markedness as fitness, i.e. the less marked a parameter-setting is, the fitter it is (see Clark and Roberts (1993) for an implementation of this idea), then we can

immediately see why syntactic change follows the S-curve. Yang (2002: 375) develops a similar idea, proposing as a corollary to his model of change that '[o]nce a grammar is on the rise . . . it is unstoppable'.

Niyogi and Berwick (1995; 1997) show that it is possible to model the interaction of a set of grammars, a learning algorithm of a particular type, and a quantity P_i 'the distribution with which sentences of the i th grammar g_i . . . are presented if there is a speaker of g_i in the adult population' (1995: 1) as a **dynamical system**. (Lass (1997: 293–304) gives a very useful introduction to some of the concepts of dynamical-systems theory in the context of historical linguistics; see note 32 below.) We can think of P_i as relating to the nature of the PLD. They apply their model to the case of the loss of V2 in the history of French (see §1.3.2). They demonstrate that if the initial condition assumes a homogeneous speech community (i.e. that P_i is uniform), the V2 system dies out rather slowly ('within 20 generations, 15 percent of the speakers have lost Verb second completely' (1995: 4)). Very interestingly, they go on to show that if the initial condition involves a mix of V2 and non-V2 grammars, V2 is lost across the population much more quickly. They say that '[s]urprisingly small proportions of Modern French [i.e. a non-V2 grammar – IGR] cause a disproportionate number of speakers to lose Verb second' (1995: 6). It seems, then, that variation might lead to change simply because of the ways in which the alternate grammars, the learning algorithm, and the distribution of PLD interact dynamically. This conclusion is confirmed in Niyogi and Berwick (1997), where, alongside the conclusion just described, that V2 becomes unstable if only a very small proportion of non-V2 examples are introduced into the sentences presented to the learning algorithm, they show that null subjects are not lost under these conditions, and so the system does not replicate the actual historical development of French. An important further result is that they are able to derive the S-curve as the typical (but not the only) form of propagation of a new system through a population. These simulations are very suggestive and arguably represent, as Niyogi and Berwick (1997: 16) point out, a first real effort to 'place historical linguistics . . . on a scientific platform.'

Niyogi (2004) develops these ideas further, and shows in detail how learning theory can contribute to our understanding of language change (as well as relating it to other questions, such as the evolution of language), once the fundamental concepts are properly formalized. As in the case of Niyogi and Berwick (1995; 1997), the central idea is that, given a learning

algorithm, a probability distribution of linguistic tokens across a population, and a restricted class of grammars from which to select, variability will readily result as long as the time allowed for the selection of hypotheses is restricted. In other words, the existence of a critical period for language acquisition may be sufficient to guarantee variation in a speech community after a single generation. To illustrate this idea, Niyogi (2004: 19) invites us to consider a world in which there are just two languages, L_{h1} and L_{h2} . Given a completely homogeneous community where all adults speak L_{h1} , and an infinite number of sentences in the PLD, the child will always be able to apply a learning algorithm to converge on the language of the adults, and change will never take place. On the other hand, Niyogi continues:

Now consider the possibility that the child is not exposed to an infinite number of sentences but only to a finite number N after which it matures and its language crystallizes. Whatever grammatical hypothesis the child has after N sentences, it retains for the rest of its life. Under such a setting, if N is large enough, it might be the case that most children learn L_{h1} , but a small proportion ϵ end up acquiring L_{h2} . In one generation, a completely homogeneous community has lost its pure [footnote omitted – IGR] character.

What happens ‘next’, as it were, i.e. in the third and subsequent generations, depends on the nature of the grammars of L_{h1} and L_{h2} , the quantity N , and the nature of the learning algorithm. All these issues are discussed, illustrated, and formalized in detail by Niyogi. Given the heterogeneity in any speech community, the random distribution of PLD, and the limited time for learning, change is inevitable. Moreover, ‘the dynamics of language evolution are typically non-linear’ (Niyogi 2004: 2); ‘much like phase transitions in physics, ... the continuous drift of such frequency effects could lead to discontinuous changes in the stability of languages over time’ (Niyogi 2004: 3). Modelling the interaction of learners, data, and grammars at the level of populations ‘provides some understanding of how a major transition in the linguistic behavior of a community may come about as a result of a minor drift in usage frequencies provided those frequencies pass a *critical threshold*’ (Niyogi (2004: 46), emphasis in original). In this work, we see several important strands of thought (natural languages as formal systems, the heterogeneity of speech communities, the critical period for language acquisition, and the existence of an innate capacity for acquiring grammars) combine to give us an understanding, expressible in formal, mathematical language, of how and why languages change.

In fact, we arrive at a diachronic criterion of adequacy for grammatical theories, since '[t]he class of grammars *G* (along with a proposed learning algorithm *A*) can be reduced to a dynamical system whose evolution must be consistent with that of the true evolution of human languages (as reconstructed from the historical data)' (Niyogi 2004: 240–1).

The approach just described may be workable for modelling grammar competition/coexistence in the speech community. However, an important aspect of Kroch's thinking involves the possibility that a single individual may have more than one grammar. The following quotation summarizes Kroch's point of view on this point:

One difficulty with the Niyogi and Berwick model ... is that it presumes that the competing parameter settings are located in different speakers, so that the quantitative element in syntactic change is located in the population, not in the individual. However, the data from the empirical studies that reveal the gradual nature of change are not consistent with Niyogi and Berwick's model in this respect. On the contrary, ... the variation in usage that reflects different parameter settings is found within texts ... To model this variation, it is necessary to allow for syntactic diglossia within individual authors as the normal situation during a period of change. (Kroch 2000: 722)

What Kroch refers to as 'syntactic diglossia' here means that individuals have competence in more than one syntactic system. Although the phonologies and lexica associated with each system are the same or extremely similar, the two systems differ in the value of at least one parameter and are distinct grammars by definition. (Of course, if parameter values are specified in lexical entries, then that amounts to a difference between the two lexical entries in question.)

In fact, Niyogi (2004: 333ff.) devotes some space to modelling the kind of dynamical system that arises if learners are assumed to be capable of acquiring more than one grammar. The model is applied to the changes involving the loss of V2 and null subjects in the history of French discussed in Clark and Roberts (1993) and Niyogi and Berwick (1995; 1997) (see §1.1.2 and §1.3.2.2 for more details), with the very interesting result that V2 must have been lost before null subjects were, and that the loss of V2 must have been triggered by an increase in the use of pronominal subjects (Niyogi 2004: 344–50). Both of these empirical claims are in fact probably correct; see Roberts (1993a) and especially Vance (1997) for more details. In general, then, it seems that modelling the interaction of grammars, possibly multilingual learners, and populations as dynamic systems

may be capable of contributing much to our understanding of the nature of language change.

Let us examine Kroch's concept of 'syntactic diglossia' more closely. The possibility that a single individual may have two distinct grammars must of course be acknowledged for the case of true bilinguals, individuals who have native competence in what we clearly would consider two different languages (for example, English and Italian, where at least the value of the null-subject parameter would be distinct). In many speech communities, the two grammars are in a relation of diglossia, in that one system is appropriate for a given range of sociolinguistic functions and the other for a quite distinct range. Diglossic situations typically involve a contrast between a 'high' variety, appropriate to relatively formal situations, contrasting with a 'low' variety, appropriate to more informal situations (see Ferguson (1959); Martin-Jones (2003); WLH (163)). Kroch suggests that syntactic diglossia may, where the innovative grammar is 'more native' than the conservative, prestige variety 'acquired a bit later in life' (723), account for the gradual spread of changes:

it could easily be the case that the forms in competition in syntactic diglossia represent an opposition between an innovative vernacular and a conservative literary language. Since the former would have both a psycholinguistic advantage and the advantage of numbers, it should win out over time, even in written texts. Under this model, the gradualism found in texts might not reflect any basic mechanism of language change, but rather the psycho- and sociolinguistics of bilingualism.

(Kroch 2000: 723)

This seems to imply that the 'psycho- and sociolinguistics of bilingualism', rather than facts about population dynamics or dynamic systems more generally, may underlie the S-shaped trajectory of change.

If Kroch's view is correct, we would expect to find similarities between syntactic diglossia and other forms of diglossia or bilingualism. This is entirely possible in principle, and certainly constitutes a way of understanding syntactic change which could reconcile our 'internalist' approach with the need to account for the spread of change through a speech community. Unfortunately, however, there is in practice rather little evidence in favour of syntactic diglossia in some of the cases where competing grammars have been proposed. (This does not mean, of course, that the evidence has not been lost, or was simply never recorded.)

The first problem is that, although we must postulate the use of multiple grammars by a single author in a single text, there is no clear evidence that

the authors in question control the variants in the typical diglossic way, using one grammar for formal, institutional contexts and the other in less formal contexts. For example, Pintzuk (2002: 282) treats the following two OE examples as representatives of competing grammars, since (17a) has AuxOV order and (17b) OVAux order:

- (17) a. He ne **mæg his agene aberan.**
 he NEG can his own support
 'He cannot support his own.'
 (CP 52.2)
- b. hu he **his agene unðeawas ongietan wille**
 how he his own faults perceive will
 'how he will perceive his own faults'
 (CP 22.21–2)

Pintzuk argues that (17a) is generated by grammar with head-initial TP (IP in her terminology), since auxiliaries are taken to occupy T, while (17b) is generated by a grammar with head-final TP. Both examples are taken from a single text, Alfred's translation of Gregory the Great's *Cura Pastoralis* written in the late ninth century. So, given Pintzuk's analysis, the sentences in (17) support Kroch's assertion that competing grammars may be found in a single text by a single author. However, they also raise a problem for the notion of diglossia: it is not clear whether one would want to categorize Alfred's style as 'high' or 'low'. The translations of various Latin works such as the *Cura Pastoralis* were carried out on Alfred's orders to preserve knowledge of Latin culture in England at a time of a decline in learning owing to the Danish invasions: see Mitchell and Robinson (1992: 204),¹⁸ but there seems to be no reason to suppose that the translator is consciously manipulating 'high' or 'low' registers in using these different word orders. Certainly, no such claim is made by Pintzuk, although she comments (278) that 'competition occurs within an individual and can be understood as code-switching or register-switching'. The notion that the different word orders reflect different grammars in a diglossic relation does not receive any further discussion or support, however.¹⁹

¹⁸ These translations of Latin works are sometimes referred to as 'vernacular', for example, by Mitchell and Robinson (*ibid.*), but in this context this simply means OE as opposed to Latin.

¹⁹ Harris and Campbell (1995: 86) also point out, discussing Lightfoot's (1991: 136–7) appeal to diglossia in a similar context, that this line of reasoning is likely to lead to the postulation of 'a plethora of grammars'.

Pintzuk's mention of code-switching raises another possibility. It seems clear from studies of code-switching that a single speaker may employ two distinct grammars even in a single sentence.²⁰ The following example of English–Spanish code-mixing illustrates this:

- (18) No creo que son fifty dollar suede ones.
 '(I) don't think that (they) are ...'
 (Poplack 1980: 598, cited in Muysken (2000: 261))

Here we see the main clause and most of the embedded clause, as far as the post-copular nominal predicate (*fifty dollar suede ones*), is Spanish. Being Spanish, null subjects are allowed: the subject of each clause is null. So in code-mixing, speakers can produce sentences parts of which are generated by a grammar with one parameter setting, and parts by another grammar with an opposite parameter setting. One might think that (18) is a Spanish sentence with English words inserted in it, but if parameters are specified by lexical entries, then, although the negative value of the null-subject parameter is not expressed by the English portion of (18), there is a switch in grammars along with the switch in lexical items; this can in fact be seen from the order of prenominal modifiers in the DP *fifty dollar suede ones*, which shows an order of adnominal modifiers that is impossible in Spanish.

However one accounts for the phenomenon illustrated in (18) (and see Muysken (2000: 261ff.) for an interesting discussion), it shows that different grammars may be used even in a single sentence. If this is true for code-mixing in general, and if the 'syntactic diglossia' discussed by Kroch is also subject to code-mixing, then we expect to find it at the subsentential level. But this creates an empirical problem, at least for Pintzuk's study of word-order change in the history of English. It is well-known that OE allowed a wide range of subordinate-clause word orders (see the discussions in §1.6.2 and §2.5). In fact all the logically possible orderings of auxiliary (Aux), verb (V), and direct object (O) are attested, as pointed out by Pintzuk (2002: 282; 2005: 252), with the striking exception of the order VOAux (Pintzuk (2002: 282, n.6; 2005: 253); see also Kiparsky (1996: 162); Roberts (1997: 416)).²¹ But this

²⁰ Muysken (2000: 4) prefers the term 'code-mixing', suggesting that 'code-switching' already implies a particular analysis of the processes involved.

²¹ This order is also unattested at any stage of the history of Icelandic, despite all the other possibilities being found (Hróarsdóttir (1999: 206); Rögnvaldsson (1996: 73)). It is also absent in Modern Dutch and German. Den Besten (1986) suggested that this might in fact be universally true. Certainly, to my knowledge, no clear case of this order has been attested. The orders AuxOV and OVAux are illustrated in (17a, b) above. VAuxO and AuxVO are illustrated in (133a, b) and OAuxV in (134b) of Chapter 1.

order should be allowed if subsentential code-switching is possible where the grammars are in competition, and if head-initial and head-final TPs are in competition. That is, a structure such as the following should be possible (extrapolating away from vP, which is not assumed by Pintzuk):²²

(19) [_{TP} DP-subject [_{VP} V DP-object] [_T Aux]]

Muysken (2000: 261–2) gives evidence for linearly discontinuous code-mixing of the type that would produce the structure in (19). Here TP is produced by a head-final system and VP by a head-initial one. So such orders are predicted by Pintzuk's approach, and yet are not found; therefore they must be ruled out by stipulation (see Pintzuk (2005: 253)). This point is also made by Fuß and Trips (2002: 183), who propose a refinement of the competing grammars model in order to deal with it.²³

So it seems that the competing grammars must be allowed to coexist in a single individual, without standing in a diglossic relationship and without the kind of possibility of subsentential code-mixing that is common in bilingual individuals and speech communities. Muysken (2000: 1–2) points out that a 'growing number of studies have demonstrated ... that many bilinguals will produce mixed sentences in ordinary conversations ... for some speakers it is the unmarked code in certain circumstances'. For speakers commanding two very similar systems, which are all but identical lexically and phonologically and minimally different in some syntactic parameter, as is proposed for OE by Pintzuk for example, such code-mixing should be highly prevalent. But the absence of the VOAux order suggests that this is not the case, or at the very least poses a serious analytical problem for accounts like those put forward by Pintzuk. So the first problem with the competing grammars approach is that it predicts unattested structures, given the expected effects of code-mixing.

²² In Pintzuk (2005: 253, (5)), a double-VP structure is assumed, with the VPs labelled VP₁ and VP₂ and the auxiliary moving from the higher V, V₁, to T (her I). Presumably, VP₁ could be assimilated to vP.

²³ Fuß and Trips propose that only lexical categories can vary for the head parameter, while functional categories are universally head-initial. If 'massive movement' of the type described in §2.5.4 is not allowed, then this is enough to rule out the VOAux order, since T will always take its VP complement to its right. At the same time, the orders attested in OE are allowed. Clearly, however, one can question the theoretical basis of this directionality distinction between lexical and functional categories. Moreover, a number of examples of Greenbergian cross-categorical harmony of the type illustrated in §1.6.1 will be hard to account for, for example, the order NP-D, or clause-final complementizers. (See the discussion following the Malayalam examples in (130) of §1.6.1.)

The second problem, which has often been pointed out, is that the competing-grammars hypothesis creates learnability problems. Kroch (1994: 184) argues that ‘the learner will postulate competing grammars only when languages give evidence of the simultaneous use of incompatible forms’. Similarly, Niyogi (2004: 336) states that the fact that ‘learners do not attain a single grammar ... reflects the fact that a single grammar is not adequate to account for the conflicting data they receive’. This must be true in the case of ‘true’ bilingualism: presumably a child exposed to PLD from both Italian and English is able to tell, on the basis of a host of cues including of course phonological and lexical evidence, that s/he is confronted with two systems. But this is not so clear if the two systems are minimally syntactically distinct, not obviously sociolinguistically distinct (since they coexist in the same speaker and the same text) and not, or only minimally, distinct in terms of the lexicon and phonology. Kroch (2000: 722) states that ‘[h]ow learners acquire diglossic competence is, of course, an important issue for language acquisition, but there is no doubt that they do’. Niyogi (2004: 336) proposes a way to formally model how learners may estimate the instances of different grammars in the PLD, but this does not really give rise to diglossia in the sociolinguistic sense, and it is not clear under what conditions multiple grammars would be postulated rather than a single one. Hence the acquisition of ‘covert diglossia’ may be a genuine problem, or in any case it raises a complication which is not raised if we do not postulate competing grammars, and hence the complication is justified to the extent that competing grammars play a role in explaining syntactic change.

If something like the general format for parameters introduced in (54) of Chapter 3 is right, then each parameter is associated with a marked and unmarked value. As we pointed out, if the marked value is not triggered (i.e. expressed in the PLD), the unmarked value is assumed. This view is not straightforwardly compatible with the idea of competing grammars. Let us try to put the problem in general terms. Suppose we have a parameter p with two values m and u , the marked and unmarked ones. (Recall that markedness may be determined in relation to other aspects of the system; this does not alter the illustration here.) Then, if the cue or expression for value m of p is present, p will be set to that value; otherwise (including in the case where there is absolutely no evidence relevant to p in the PLD, owing to the values of superordinate parameters – cf. the discussion of Baker’s ‘periodic table’ in (60) of Chapter 3) it will be set to u . This approach appears to leave little room for the postulation of competing grammars:

‘evidence of the simultaneous use of incompatible forms’ would presumably favour the postulation of the marked value, since this would effectively constitute evidence for this value, given that the default value by assumption requires no evidence. In all and only cases where the marked value is not expressed in the PLD, the unmarked form results.

Moreover, even if we weaken the approach described in Chapter 3 and allow for the possibility that ‘the simultaneous use of incompatible forms’ may allow the postulation of two systems in competition (we will see an interesting suggestion to this effect, again due to Kroch, below), questions nevertheless arise as to how such a system might change: why would a later generation choose just one of the values if it is possible – in terms of the markedness relation between the values and their expression in the PLD – to choose both? We could see this as a particularly tricky case of the Regress Problem. Language contact may perturb the PLD in such a way as to produce the required outcome by reducing the robustness of expression of the marked value (through the effect of the Alien Corpus; see (23) of Chapter 3); we will discuss this possibility in more detail in §5.2. It should also be borne in mind that ‘our grasp of the relationship between the evidence presented to a learner and the grammar acquired is still imprecise’ (Kroch 2000: 700). Nevertheless, we can see that the postulation of competing grammars, although a possible example of what WLH called ordered heterogeneity, may not be without problems for the account of setting and changing parameter values that we put forward at the end of Chapter 3.

Leaving aside language contact for now, there are three ways in which competing grammars could be acquired in terms of the approach to parameter-setting we have put forward. The first possibility is that the syntactically distinct systems are also distinct in some other ways: if there are clear phonological and lexical differences systematically associated with the parametrically differing systems, then presumably acquirers will keep the two systems apart. Again, this is what must happen in the case of ‘true’ bilingualism.²⁴ By assumption, this is not what happens in the case of competing grammars.

²⁴ There is evidence from first-language acquisition that both phonological and lexical acquisition begin in the first year of life. In particular, very young children (in the first month of life) seem able to distinguish their native language from other languages, and begin to be able to discriminate consonants. These findings are documented in detail in Guasti (2002, Chapter 2); see in particular her Table 2.1, 25. It is possible, then, that two completely distinct systems are kept apart from a stage prior to the onset of the acquisition of syntax.

The second possibility is to postulate that children associate the competing grammars with some sociolinguistic or contextual marker. As Denison (2003: 58) and Croft (2000: 185–8) point out, this can also account for change, and possibly for the S-shaped propagation of change: if an innovating form has a social value, this may favour it in the grammar competition. The idea of distinct systems being available to speakers in a given speech community is discussed in WLH (161–4). They insist that the two systems must be functionally distinct (162) for two reasons. First, in order to be of interest for the theory, the two systems must be in competition. Second, there must be ‘a rigorous description of the conditions which govern the alternation of the two systems. Rules of this sort must include extralinguistic factors as governing environments’ (161–2). These extralinguistic factors include the kinds of factors which have been much discussed by sociolinguists: age, sex, gender, social class, ethnicity, etc. What has not been clearly shown in the studies of grammars in competition following on from Kroch (1989) is that the putatively competing grammars are functionally distinct in WLH’s sense. Indeed, if anything, the evidence suggests that they are not, as examples like (14) indicate (although of course it must always be kept in mind that the relevant evidence is extremely flimsy).

The third and arguably most interesting way to think of the acquisition of competing grammars would be to invoke, presumably at the level of UG, a ban on true doublets. As Pintzuk (2003: 525) suggests, one could, following Kroch (1994: 180), assume that the presence of ‘syntactic doublets in the lexicon’ will give rise to variation and change. As she says, Kroch (1994) ‘suggests that in syntax, as in morphology, doublets that are semantically and functionally non-distinct are disallowed; and that doublets of this type ... compete in usage until one of the forms wins out’ (Pintzuk 2003: 525). In morphology, the Blocking Effect, defined as ‘the non-occurrence of one form due to the simple existence of another’ (Aronoff (1976: 43), bans doublets, hence the non-existence of **graciousity* is due to the existence of *graciousness*. In other words, acquirers could be led to the postulation of two systems by the mere existence of syntactic doublets. This could follow if system-internal optionality were not allowed by UG. Then the existence of doublets, analysable in principle as options, would have to be seen by acquirers as contradictory properties, and hence as distinct systems. Such a line of reasoning could clearly be applied in the case, for example, of EPP features of functional heads. Moreover, if these are features of lexical

entries, one might think that this is no different in principle from something as simple as the variant pronunciations of *economics* (with initial /i:/ or initial /ɛ/), with the exception that free variation is simply not allowed.

This approach is very appealing, and technically compatible with earlier versions of minimalism such as Chomsky (1993; 1995) where, as we mentioned in the previous section, optionality was not allowed. However, as we have seen, syntactic doublets genuinely do appear to exist: the Russian example in (11) illustrates this. More generally, true optionality, with no functional or semantic import, does seem to be allowed in the current system, as we saw in the previous section. If that is true, then Kroch's (1995) 'no-doublets' assumption is empirically wrong, and therefore this way of accounting for the acquisition of competing grammars is not available. I will return to the possible implications of true optionality below.

I conclude, somewhat reluctantly, that the competing grammars idea, although in principle providing a way of accounting for the spread of change through a speech community and over time in a way which would be broadly compatible with the approach to syntactic change we described in the previous chapter, has not really been shown to be free of the twin problems of allowing for unattested examples through subsentential code-mixing and of posing the learnability problem of when such a system should be postulated. A clear demonstration of the social value of one or the other competing grammars would, however, illustrate the utility of this approach (and we will see a likely example of this in our discussion of Brazilian Portuguese in §4.2.6 below). And, of course, we cannot exclude the possibility that the competing grammars postulated by Kroch, Pintzuk, and others for the earlier stages of various languages for which we have little or no sociolinguistic information, did have a social value whose nature has been completely obscured by the passage of time and the nature of the extant texts.

4.2.4. *Formal optionality again*

WLH (166–76) discuss the notion of system-internal variation. We have seen that the current model of syntax allows for this (cf. the discussion of (10) in the previous section). So, a further possibility for accounting for the spread of change lies in true optionality. Let us reconsider the example of

optionality of pied-piping in *wh*-movement constructions in Russian from Biberauer and Richards (2006):

- (11) a. Č'ju knigu ty čital?
 whose book you read
 b. Č'ju ty čital knigu?
 whose you read book
 'Whose book did you read?'

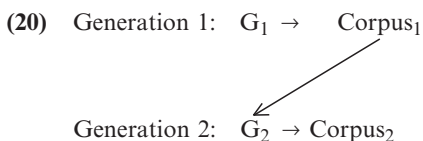
(Biberauer and Richards 2006, (28), 21)

According to Biberauer and Richards, the difference between Russian and English is that in Russian *wh*-words are DPs, while in English they are Ds. Both languages have overt *wh*-movement in interrogatives, i.e. in both languages a [+*wh*] C forces movement of a *wh*-XP to its Specifier in virtue of having an EPP feature. Russian therefore automatically has an option lacking in English: that of moving the *wh*-word alone to SpecCP. Russian thus provides a simple example of a single grammatical system, with all parameters set to determinate values, which can nevertheless produce doublets, i.e. true optionality where there is no semantic, functional, or social difference. It might then suffice for one of the doublets to take on social value in order for one of the variants to become predominant (in certain social groups) and hence for a change to take place. However, to fully understand change in this way, we need to bear in mind the interaction of structural, sociolinguistic, and psycholinguistic factors. The structural factor is that UG disallows the possibility of non-XPs undergoing *wh*-movement for reasons that are unclear, but the observation that no language only has the option instantiated by (11b) appears to be correct (see Gavarrò and Solà (2004: 10), who observe that this is also true for child language). Hence, this doublet cannot oust the other. On the other hand, a change of the kind whereby option (11b) disappears and only (11a) remains (which we suggested, following Biberauer and Richards, may have taken place in the history of Greek) is favoured by markedness considerations of two kinds: first, reanalysis of *wh*-elements from DP to D represents a structural simplification (in the sense of (22) of Chapter 3); second, the elimination of (11b) and the retention of (11a) involves a shift from a superset to a subset grammar, and hence is favoured by the Subset Principle (see §3.4.3). But we can consider that the taking on of social value on the part of the variant in (11a) may have had the effect of skewing the PLD for speakers in such a way that the unmarked option of the grammar which only allows (11a) was preferred. Here we see how sociolinguistic factors

may interact with other factors as a way of solving the Regress Problem. It may be that such an interaction of sociolinguistic, psycholinguistic, and structural factors is required in order to have a complete picture of how syntactic change works. (Of course, the account of this particular change in the history of Greek remains very sketchy owing to the lack of relevant data.) However, if Niyogi and Berwick's (1995; 1997) results, as described above, are borne in mind, it may be enough to simply have variation in the PLD; the social value of the variants may influence the change in various important ways, but change is guaranteed by the mere presence of variation. For this kind of scenario to have any generality, we need to formulate parameters in such a way that true syntactic doublets may be a recurrent possibility; this has not been done in most of the work on parameters in recent years, and an important research question concerns the extent to which it can be done. As Biberauer and Richards (2006) observe, however, the technical make-up of the current system leads us to expect formal optionality: '[o]ptionality ... would have to be stipulated *not* to exist'.

4.2.5. *Abduction and actuation*

A final question brings us back to matter of the actuation of change. We have offered a purely 'internalist' account of this, seeing change as driven by abductive reanalysis, associated with parametric change, in language acquisition. Let us consider again the schema for abductive reanalysis that we discussed in §2.1.1:



In our earlier discussion, we referred to G_1 as the 'parental grammar' and G_2 as the children's grammar. Corpus_1 was seen as the output of the parental grammar. We observed that this description involved a certain amount of idealization.

In fact, WLH (145–6) criticize Halle's (1962) reanalytical model of phonological change, which is quite similar to what is schematized in

(20), on several grounds, two of which are relevant here. First, they point out that the speech of older peers, only slightly older than the acquirer, plays a more important role in determining the nature of the acquired system than does parental speech. (This point is also taken up by Croft (2000: 48).) Second, they observe that change is not purely intergenerational, but that changes appear to last longer, and continue in the same direction, across generations.

The second point concerns gradualness: as we have seen, grammatical and sociolinguistic factors can cushion the effects of parametric change and cause it to appear to be temporally diffuse; at the level of the speech community this is very likely if the community has any sociolinguistic complexity, as all speech communities do. What it implies is that successive generations can replicate the ordered heterogeneity of the speech community, whether this is done through the postulation of competing grammars or through the exploitation of formal options in the grammatical system, or a combination of the two. In other words, language acquirers can acquire variation. This point is also relevant to the discussion of drift in the next section.

The first point is well-taken, and requires us to see the PLD as more diverse than we have hitherto been assuming. Again, this means that orderly differentiation must be present in the PLD and that acquirers must be sensitive to it. Rather than referring to G_1 in (20) as the parental grammar, then, we should call it the system (or systems) underlying the PLD. That this may not be produced by parents, but rather by older siblings or peers, does not in itself require that we abandon the concept of abductive change. And of course, we should not refer to G_1 and G_2 as properties of different generations (itself a highly diffuse concept, as WLH (114) point out), but rather as properties of a relatively older group providing PLD for a relatively younger group. This implies that where there is abductive reanalysis in G_2 of (20), the age discontinuity between the two groups may be rather small, perhaps less than ten years. This may be a further source of the observed gradualness of change. However, we are able to retain the notion of discrete parameter values underlying the formal distinctions between the grammatical systems.

WLH (184–5) present the following summary account of their proposals for how change spreads (the ‘transition problem’ in their terminology):

This transition or transfer of features from one speaker to another appears to take place through the medium of bidialectal speakers, or more generally, speakers with

heterogeneous systems characterized by orderly differentiation. Change takes place (1) as a speaker learns an alternate form, (2) during the time that the two forms exist in contact within his competence, and (3) when one of the forms becomes obsolete.

This approach can be taken over into the parameter-based account of syntactic change being proposed here, with the ‘alternate forms’ in question being either options generated by a single system or competing grammars. The new parametric option (represented either by a new grammar or a resetting of a parameter such that a new option is introduced) enters the speech community and some individual’s competence at stage (1), and the old parametric option (represented either by an old grammar or by a resetting of a parameter such that an option is lost) disappears at stage (3). At stage (2), one option replaces the other as it spreads through the community, with a changing proportion of individuals having two grammars or a setting of the relevant parameter permitting formal optionality. Clearly, this scenario gives rise to more apparent gradualness. It can also account for the typical S-curve of change as one grammar replaces the other through the community at stage (2). Moreover, it is broadly in accordance with the results of Niyogi and Berwick’s (1995; 1997) dynamic models as briefly summarized above. Thus we have a general model for gradual spread of change through a speech community, which is compatible with the results of sociolinguistic research on synchronic variation of the kind reported above by Labov (1966; 1972; 1994; 2001), and which nevertheless allows us to maintain an acquisition-driven view of the actuation of change and a notion of discrete, binary parameters as the locus of variation among grammatical systems. Of course, much more needs to be done to really establish this view, and much more direct evidence for it needs to be gathered, but it represents, in my opinion, a promising research programme for getting a full picture of the nature of syntactic change.

4.2.6. *Change in progress? Null subjects in Brazilian Portuguese*

An example of ongoing change may come from Brazilian Portuguese (BP). This variety appears to be in the process of losing fully productive null subjects, i.e. changing the value of the null-subject parameter. European Portuguese (EP) shows all the standard hallmarks of a null-subject language (see §1.1.1):

- (21) a. *Null definite pronominal subjects:*
 Telefonaram ontem.
 called-3pl yesterday
 ‘They called yesterday.’
- b. *‘Free inversion’:*
 Telefonou ontem o João.
 called-3sg yesterday John
 ‘John called yesterday.’
- c. *Wh-movement of a subject over a finite complementizer:*
 Que aluno disseste que – comprou um computador?
 which student said-2pl that – bought-3sg a computer
 ‘Which student did you say bought a computer?’
 (Barbosa, Duarte, and Kato 2005: 1)

BP, on the other hand, does not allow free inversion, except with unaccusative verbs – see Kato (2000).²⁵ It does, however, allow null subjects as in (21a) and wh-movement as in (21c), but the incidence of 3rd-person null subjects compared to EP is considerably reduced. Barbosa, Duarte, and Kato (2005: 7) report the results shown in Table 4.1 for the rate of overt and null 3rd-person subjects in the two varieties.

Moreover, BP allows overt inanimate subjects much more readily than EP, and allows an overt embedded pronoun subject to corefer with a matrix DP (cf. the discussion of (15)–(16) in §1.1.1). Both of these properties are illustrated in (22) (underlining indicates intended coreference):

- (22) A casa virou um filme quando ela teve de ir abaixo.
 the house became a movie when it had to go down
 ‘The house became a movie when it had to be pulled down.’

Barbosa, Duarte, and Kato illustrate a host of other respects in which EP patterns like a canonical null-subject language and BP diverges; these

Table 4.1 Occurrences of null and overt subjects in EP and BP (Barbosa, Duarte, and Kato 2005: 7)

Variety	Null subject	Overt subject	Total
EP	126 (78%)	36 (22%)	162 (100%)
BP	63 (44%)	79 (56%)	142 (100%)

²⁵ According to Ilza Ribeiro (p.c.), inversion of the subject of an unergative intransitive as in (21b) is possible in BP, with a very strong focus on the inverted subject.

include subject left-dislocation with an overt resumptive subject pronoun in BP, postverbal emphatic subject pronouns in EP, the possibility of interpreting overt pronouns as bound pronouns in BP, and the possibility of relative-clause exposition in BP. The contrasts regarding left-dislocation are illustrated in (23):²⁶

- (23) a. A Clarinha, ela cozinha que é uma maravilha. (BP)
 the Clarinha, she cooks that is a wonder
 b. A Clarinha, __ cozinha que é uma maravilha. (EP)
 the Clarinha, cooks that is a wonder
 ‘Clarinha cooks wonderfully.’
 (Barbosa, Duarte, and Kato 2005: 3–5)

BP has also lost other properties that have been linked to the positive value of the null-subject parameter, namely clitic-climbing and enclisis. (See (29) and (30) of §1.2.1; the absence of these features in BP is discussed in various papers in Roberts and Kato (1993).)

So we see that contemporary spoken BP is not a null-subject language, at least not in the ‘standard’ way that EP is (along with Spanish, Italian, Greek, etc.).

Duarte (1993; 1995) documents the use of overt subject pronouns in popular plays and television shows from the first half of the nineteenth century up to 1992. Her results, shown in Figure 4.4, indicate that the rate of pronominal subjects rose from 20 per cent in the mid-nineteenth century to 74 per cent by the end of the twentieth century.

It is possible to discern an S-curve here, giving the result that contemporary BP is coming close to completion of this change.

Duarte suggests that the change was caused by a reorganization of the pronominal system in such a way that formerly 3rd-person DPs came to be used as 2sg, 2pl, and 1pl forms, replacing the earlier pronouns, with the result that 3sg verbal inflection is now used in these persons as well. The effect is a levelling of the verb paradigm, as shown in (24):²⁷

²⁶ The authors point out that the string corresponding to the BP one in (23b) is grammatical in EP, but it carries a sense of redundancy and is ‘uttered only when the speaker is hesitating’ (6), which is not the case in BP.

²⁷ The replacement of the 2sg familiar pronoun *tu* by the polite form *você* is reminiscent of the replacement of the 2pl intimate pronoun *thou* by the 2pl/formal *you* in ENE (see Lass (1999: 148–54) and the references given there). It is not quite the same, however, in that *você* is etymologically 3rd-person; the original 2pl form in Portuguese being *vos*.

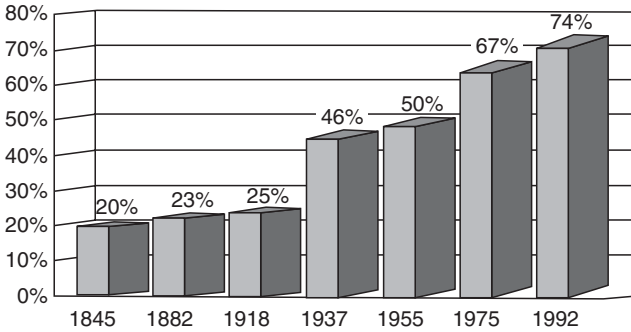


Figure 4.4 The rate of overt pronominal subjects in the nineteenth and twentieth centuries

Source: From Barbosa, Duarte, and Kato (2005: 3) and adapted from Duarte 1993: 112.

(24) a. *EP and former BP paradigm:*

(eu) falo	(nos) falamos
(tu) falas	(vos) falais
(ele/ela) fala	(eles) falam

b. *Reorganized colloquial BP paradigm:*

eu falo	a gente fala
você fala	vocês falam
ele/ela fala	eles falam
'I speak, etc.'	

The effect of the reorganization of the pronominal system is a reduction of the number of distinctions in the verbal inflection from six to three. As such the new system arguably falls below the threshold for 'rich' agreement (see §1.2.1) and thus, following the statement of the null-subject parameter in (54A) of the previous chapter can no longer trigger a positive setting of the null-subject parameter (see Duarte (1993; 1995) for this idea).

In these terms, we could see Stage (1) of the change in WLH's terms – the innovation of an alternate form – as the initial reorganization of the pronoun system. Stage (2) is documented by Duarte's data given in Table 4.1; thus individuals in the BP speech community either have competing grammars or the system permits formal optionality. (The former is favoured by the fact that the null-subject and non-null-subject varieties have clear social value in Brazil; see Tarallo (1993) and the papers in Tarallo (1989).) Contemporary BP is arguably not yet at Stage (3). Moreover, it is unclear that Stage (3) is really to be seen as equivalent to a totally *non*-null-subject system like English or Standard French, especially

as null subjects are fairly readily licensed in embedded clauses where there is coreference with the matrix subject, as in:

- (25) Ela costumava sentar-se em cima da cama com seu tricot,
 She used-to sit-self on top of-the bed with her knitting
 enquanto dava lições a um de nós.
 while gave-3sg lessons to one of us
 ‘She used to sit on the bed with her knitting while she taught one of us.’
 (Barbosa, Duarte, and Kato 2005: 8)

This kind of system, where null subjects are disfavoured in main clauses, but allowed in embedded clauses under coreference with a main-clause element, has been referred to as a ‘partial null-subject system’ by Holmberg *et al.* (2005). In addition to BP, they mention Finnish, Hebrew, Marathi, and (with some provisos) Icelandic, as other languages showing this kind of system. Another property that these languages share, as well as BP, is the existence of a null subject with a generic interpretation. It may be, then, that this is the kind of null-subject system BP is developing into.

The indications are, therefore, that BP is changing, at least in that it is losing its earlier status, retained in EP, as a fully-fledged null-subject language of the typical Romance type (with the clear exception of Modern French, of course). What BP is changing into is still unclear, but the dynamics of the change correspond broadly to the scenario outlined by WLH. Of course, this brief summary does not do justice to the full complexity of the situation in BP, either sociolinguistically or syntactically, but the example is illustrative. In particular, it shows us how we may be able to use the tools of principles-and-parameters theory combined with sociolinguistic theory in order to observe syntactic change in progress.

In the foregoing, I have been assuming that the ‘alternate form’ referred to in the above quotation from WLH arises through abductive reanalysis. Language and dialect contact is, as we mentioned in §3.3.3, another obvious source. We return to this topic in §5.2.

4.2.7. Conclusion

In this section I have tried to show how sociolinguistic and other considerations can interact with a principles-and-parameters-based syntactic system to give a full picture of how a change goes to completion. A number of questions remain open, of course, but the two central ideas are (i) that

change starts through the abductive reanalysis of PLD by individuals, and (ii) there is a period of variation, which may be associated with grammars in competition, formal optionality, diglossia, diffusion, etc., until one system finally takes over from the other. If the diglossic situation is sufficiently clear, in that the use of the varieties perceived as ‘high’ and ‘low’ is associated with quite unambiguous and stable social value, then the grammars may peacefully coexist across a number of generations. Something like this may be the situation in the German-speaking areas of Switzerland where Standard German is the ‘high’ variety and local varieties of Swiss German – radically different from Standard German in their syntax as in many other structural respects – represent the ‘low’ variety.

Our emphasis here has been on the nature of the process that gives rise to the initial change in the individual, but, in order to have a full picture of how changes may actually go to completion, we must take into account the period of variation and the nature of variation as well. We have seen that syntactic variation can take a variety of forms, not always indicating the presence of two distinct grammars in an individual or a speech community. One important task for future research is to clarify the differences among these types of syntactic variation, and investigate what implications these may have for syntactic change.

4.3. Drift: the question of the direction of change

4.3.1. Introduction

Here we take up WLH’s observation that changes seem to take place over longer periods than an approach which claims that it is caused by generation-to-generation transmission of PLD would lead one to expect. In fact, if abduction as in (20) relates not generations but older and younger peers within a single generation, then the prediction of the abductive approach is that change is still more rapid. So we have to face the question of the direction of change over periods longer than required for an abductive change of the type schematized in (20). But is it meaningful, or even possible, to think of syntactic change as having an inherent direction? The idea that language change is directional, and that languages pass through cycles, or ‘life cycles’, of changes was common in the nineteenth century: Morpurgo-Davies (1998: 86–8) identifies this kind of thinking in

the work of Bopp, Grimm, Humboldt, Schlegel, Schleicher, and Max Müller, and observes the influence of Darwinian and Hegelian thought on the last two of these (196–201). The idea that change is directional was stated very influentially, and eloquently, by Sapir (1921: 160ff.) in terms of the concept of drift. He says that ‘[l]anguage moves down time in a current of its own making. It has a drift’ (160), and goes on to assert that ‘[e]very word, every grammatical element, every locution, every sound and accent is a slowly changing configuration, molded by the invisible and impersonal drift that is the life of language. The evidence is overwhelming that this drift has a certain consistent direction’ (183). Moreover, we can infer the ‘drift’ of a language from its history: ‘[t]he linguistic drift has direction . . . [t]his direction may be inferred, in the main, from the past history of the language’ (165–6). He suggests three connected drifts in the history of English: the ‘drift toward the abolition of most case distinctions and the correlative drift toward position as an all-important grammatical method’ and ‘the drift toward the invariable word’ (180). These are all long-term drifts: ‘[e]ach of these has operated for centuries, . . . , each is almost certain to continue for centuries, possibly millennia’ (174). Each change is therefore just one in a series of changes, created by what went before and in turn creating the conditions for subsequent changes.

In this section I want to address the question of whether we could entertain a notion of ‘parametric drift’. Clearly, this question relates directly to the gradualness issue, discussed in §4.1, as well as to WLH’s criticism of the abductive model which we discussed in the previous section. But a number of further questions arise. Most prominent among these is the matter of causation: if there is such a thing as parametric drift, what causes it? We do not wish to invoke Hegel-style laws of history, or the nineteenth-century organic metaphor of languages (or grammars) as going through a ‘life cycle’. But if grammars are transmitted discontinuously from generation to generation, or from older to younger peers, through language acquisition, how can acquirers know which way the system they are acquiring is drifting? As we have seen, this point has been very forcefully made by Lightfoot (1979; 1991; 1999). A second question is: if drift does exist, what are the natural directions for it? Third, how does this concept relate to the general idea that syntactic change is subject to the Inertia Principle, as discussed in §3.2?

Here I will suggest that parametric drift can be thought of as a cascade of changes, a kind of ‘domino effect’ in the parametric system, whereby an

initial, exogenous change destabilizes the system and causes it to transit through a series of marked states until it eventually restabilizes as a relatively unmarked system again. The series of marked states could in principle cover several cohorts of acquirers, with each successive group of acquirers being led to reanalyse different aspects of the PLD which have been rendered marked by an earlier change. I will briefly illustrate this idea by looking at changes affecting the English auxiliaries in the ENE period, some of which we have already seen in earlier sections.

This approach answers the questions raised above in the following ways: markedness may be the force which underlies parametric drift and determines the natural direction of change; it is also what causes the reanalyses, explaining why the Inertia Principle does not hold in these cases (although I will tentatively suggest that the Inertia Principle does hold in a slightly different way).

4.3.2. *Typological approaches to drift*

Typological studies have often emphasized the directional nature of change. The earliest kind of structural typology proposed was the morphological typology which distinguishes **isolating, agglutinating, and inflectional languages**. Isolating languages tend to lack inflectional affixes, marking grammatical notions either by word order or by separate particles of various kinds: Chinese and Vietnamese are usually cited as good examples of this type. Agglutinating languages add affixes to roots in a predictable way, with something close to a one form–one meaning correlation, while inflectional languages add affixes to roots in a way subject to various kinds of phonological and morphological conditioning. Malayalam is a good example of an agglutinating language (see §1.6.1), and the older Indo-European languages such as Latin, Classical Greek, or Sanskrit are good examples of inflectional languages (Comrie (1989: 42–52) gives details and examples). This typology is usually attributed to the early nineteenth-century linguists Friedrich and August Schlegel (but see Morpurgo-Davies (1998: 71–5) on the seventeenth- and eighteenth-century antecedents to this typology), and formed the basis of a directional account of morphological change in Schleicher (1861–2: 4, 342–3). Schleicher proposed that languages change from isolating to agglutinating to inflectional in that direction and not in the opposite one. Change in morphological type has thus been seen as unidirectional (see Croft (2003: 252–3) on this).

As we saw in §2.5.2, the idea of long-term typological change in word order is associated above all with W. Lehmann (1973) and Vennemann (1974). Lehmann (1973: 55) proposed that languages which are not typologically consistent in terms of his OV vs. VO typology are undergoing change. He discusses directional change from OV to VO in Indo-European (55–8) and mentions the possibility of VO changing to OV in Tocharian (57). We could apply this idea to the history of English, and observe that English has been drifting from OV to VO since at least the Proto-Germanic period, and has not yet completed the drift (since English retains AN and DemN orders, at least). Vennemann's (1974) approach, based on the Natural Serialisation Principle, also runs into the difficulty that we are led to regard 'mixed' systems as persisting over very long periods. We have quoted Vennemann's (1974: 353) remark that 'a language may become fairly consistent within a type in about 5000 years' (for example, English). Both Lehmann and Vennemann proposed long-term, directional, typological drift from OV to VO as a mechanism of change.

A well-known and very striking example of apparently directional typological change comes from Greenberg's (1980) study of word-order change in the Ethiopian Semitic languages. Greenberg looked at synchronic word-order differences among a number of these languages (Ge'ez, Tigre, Tigrinya, fourteenth-century Amharic, Modern Amharic, Old Harari, and Harari) and observed that Ge'ez had free word order in the clause, with a tendency towards VSO, and has AN order alternating with NA in nominals, along with NGen (alternating with GenN) and Prepositions. Tigre is SOV and AN, but predominantly NGen and Prepositional. Tigrinya is the same, but with GenN. Amharic of both periods is like Tigrinya but with GenN and a growing tendency for Postpositions. Finally, Harari is basically postpositional. As Croft (2003: 249) says, these systems 'differ from each other in small enough ways that the actual historical process can be perceived, just as motion is perceived in a sequence of stills from a movie'. Greenberg (1980: 238–41) also compares the situation in the Ethiopic languages with what seems to be a parallel situation in the Iranian languages.

The important point here concerns directionality. Croft (2003: 250–1) observes that if we concentrate on three of the variant word-order dyads, AN/NA, GenN/NGen, and Prepositions vs. Postpositions, we can observe the following combinations:

- (26) NA and NGen and Prep
 AN and NGen and Prep
 AN and GenN and Prep
 AN and GenN and Postp

Croft says that this can be seen as a unidirectional historical process, whereby each change leads to the next and the reverse process is not found. Unidirectional changes of this kind impose ‘a major constraint on possible language changes. In fact, [they] cut[s] out half of the logically possible language changes’ (2003: 251). Discovering such unidirectional processes is therefore a major goal of diachronic typological linguistics. Here we are again dealing with long-term typological drift.

One objection which might be made here is that we could in fact add an initial stage to (26), on the basis of Greenberg’s data from Ge’ez, if we take into account clausal word order. We could, abstracting slightly away from the possibility of alternate orders, bring the OV/VO dimension in and restate (26) as (27):

- (27) VO and NA and NGen and Prep (Ge’ez)
 OV and NA and NGen and Prep (Tigre)
 OV and AN and NGen and Prep (Tigrinya)
 OV and AN and GenN and Prep (Fourteenth-century Amharic)
 OV and AN and GenN and Postp (Harari)

It seems then that the first of the series of changes, taking Ge’ez to represent the most conservative stage, was from VO to OV.²⁸ But we have remarked that this is an unusual change, and that OV to VO is more common (cf. English, North Germanic, Romance, Celtic, Greek, Slavonic, and Western Finno-Ugric, all of which have undergone the change from OV to VO at some point in their history).²⁹ Croft does not deal with this specific point,

²⁸ It is worth noting that (27) conforms to the markedness convention for the relations among word-order parameters that we tentatively proposed in §3.5 (57), since VO first changes to OV triggering the general reversal of head-complement orders.

²⁹ As we noted in §1.6.1, several Iranian languages are typologically unusual. Greenberg (1980: 240) notes that we would expect Old Persian to be VO, given the synchronic variation in the languages he documents. But it is OV. He goes on to suggest, following Friedrich (1975), that, while Proto-Iranian was SOV, Avestan allowed VSO order and showed other VO traits (prepositions and NG order). He concludes that ‘[t]his was the situation when a wave of SOV spread over Iranian territory during the middle Iranian period’ (240). It seems, then, that these languages began developing from OV to VO, but, for unclear reasons, this change was reversed.

but he does suggest that ‘language processes that appear to be bi-directional often turn out to represent two distinct unidirectional changes that involve different mechanisms of language change or involve different intermediate language states’ (2003: 251). This point is well-taken in principle, although it is not clear what distinguishes this instance of VO > OV from the apparently much commoner OV > VO we have discussed elsewhere.

4.3.3. *Drift and parametric change*

Lightfoot has criticized approaches to syntactic change which invoke typological drift as a mechanism in various places, as we have already mentioned (see §2.5.2 and §4.1). The most detailed and explicit discussion is in Lightfoot (1979: 385–99). He criticizes this general approach to word-order change on three grounds, all of them essentially stemming from the single general point that grammars must be seen as being replicated through language acquisition, rather than transmitted in any direct way across the generations (or from older to younger peers). We gave the crucial quotation in the discussion of gradualness above, and repeat it here:

Languages are learned and grammars constructed by the individuals of each generation. They do not have racial memories such that they know in some sense that their language has gradually been developing from, say, an SOV and towards an SVO type, and that it must continue along that path. After all, if there were a prescribed hierarchy of changes to be performed, how could a child, confronted with a language exactly half-way along this hierarchy, know whether the language was changing from type *x* to type *y*, or vice versa?

(Lightfoot 1979: 391)

From this he concludes: ‘[t]herefore, when one bears in mind the abductive nature of the acquisitional process, the concept of an independent diachronic universal (i.e. unrelated to the theory of grammar) becomes most implausible’. He identifies three ways in which one could countenance long-term drift against the background of ‘the abductive nature of the acquisitional process’: racial memory, ‘mystical metaconditions on linguistic families or goal-oriented clusters of changes’ (395), and typological drift as seen in the previous subsection. Of these, clearly the first can be disregarded with no further discussion. The second was proposed in Lakoff’s (1972: 192) account of the drift towards analyticity in various Indo-European languages. He objects that such metaconditions run the risk

of postulating diachronic grammars, a notion inherently incompatible with the nature of the language acquisition and, as such, untenable.³⁰

The third way in which one could countenance long-term drift is of course the kind of typological drift that we illustrated above using Croft's summary of Greenberg's work on the Ethiopic languages. Here the criticism is that there is no way of building such long-term teleology into the language-acquisition process, as the above quotation states. Moreover, Lightfoot questions the empirical basis for the diachronic generalizations: we have good long-term textual attestation of very few languages, we know that typological shifts can go in different directions (for example, OV to VO or VO to OV), and we know that even in Indo-European a fair amount of variation in basic word order can be observed: SVO (English, Romance, etc.), SOV (Indic), and VSO (Celtic). Finally, he observes that many changes do not involve alterations to observed word order: the development of the modal auxiliaries in ENE (discussed in §2.1) did not change the surface order of modals and main verbs, but was arguably nonetheless a significant change in the syntax of English. Despite all these criticisms, Lightfoot does not deny that changes can be 'provoked by earlier changes and in turn themselves provoke others' (397).

The conceptual underpinning of much of Lightfoot's critique of typological drift is fully valid: here I am advocating exactly the same general approach to syntactic change as driven through language acquisition as Lightfoot. All the points made by Lightfoot, as summarized in the previous paragraph, are well-taken, with the possible exception of the one concerning the different directions of change: here Croft's point that different directions of change might be associated with different mechanisms of change or show different intervening stages may have some merit. We can further note that VSO and SVO languages are now seen as considerably less radically different than they were in the 1970s: VSO languages are now seen as highly similar to SVO ones with V-to-T movement, with the sole difference lying in the movement of the subject to SpecTP (see §1.3.1.1).

³⁰ Such a metacondition could conceivably be a condition on grammar which is successively inherited across by the usual language-acquisition process. For Lightfoot (1979) this kind of condition would fall foul of the Transparency Principle, being very abstract. In terms of the assumptions being made here, it could derive from a markedness convention, if synthetic systems could be seen as more marked than analytic ones.

Lightfoot (1979: 396) does acknowledge, however, that ‘implicational sequences of syntactic changes do exist ... there are sets of roughly analogous changes which cluster together in independent languages’. His critique focuses on the facts that the ‘implicational sequences of syntactic changes’ are something to be explained rather than an explanatory notion in themselves. Again, this point is well-taken.

So, how can we understand ‘implicational sequences of syntactic changes’ in terms of a parametric theory of change and variation? Here, we seem at first sight to be confronted with a paradox, as noted by Roberts and Roussou (2003: 3–4). They point out that one can view the parameters of UG as defining an abstract space of variation. In that case, the natural, and arguably the usual, way of viewing synchronic variation is to see grammars as randomly scattered through this space, while the natural way to think of diachronic change would be to see it as a random ‘walk’ around this space. This is consistent with the idea that typological drift is an incoherent notion, essentially for the reasons Lightfoot gives. But, in that case, what of ‘implicational sequences of syntactic changes’ in the diachronic domain, and what of implicational relations in the synchronic domain?

Alongside the evidence for ‘implicational sequences of syntactic changes’, typological work has yielded evidence for a further type of change which strongly tends to be unidirectional: grammaticalization, the process by which new grammatical morphemes are created, either from other grammatical (or functional) elements or from lexical elements (see Croft (2003: 253–72)). In §2.2 we discussed and illustrated the approach to this phenomenon adopted by Roberts and Roussou (2003), an approach which is fully compatible with the general assumptions being made here in that grammaticalization is reduced to parameter change. Although a few isolated cases of degrammaticalization have been observed (see Roberts and Roussou (2003: 208, n. 2)), grammaticalization appears to be a pervasive phenomenon, and strongly tends to follow certain ‘pathways’; for example, minimizers or generic terms tend to become n-words and/or clausal negators, demonstratives tend to become definite determiners, and verbs of certain semantic classes tend to become auxiliaries. (Heine and Kuteva (2002) give a very extensive list of examples of this sort.) This, then, is a further example of tendential directionality in parameter change. If so, then ‘grammaticalization paths’ are a further case of implicational sequencing of syntactic change, and raise the same questions for a parametric approach as the other ones.

A further point arises from Sapir's original discussion of drift. His concern is to understand why languages differ from each other at all. He observes that social, geographical, or even individual variation do not actually provide an explanation. He says that if 'individual variations "on a flat"' were the only kind of variability in language, I believe we should be at a loss to explain why and how dialects arise, why it is that the linguistic prototype gradually breaks up into a number of mutually unintelligible languages' (1921: 160). It is in this context that he introduces drift. So, for Sapir, drift explains the very existence of different grammatical systems. Some explanation of this phenomenon is also required even if a parameterized syntax is assumed: why, even if there are different parameters, should there be different systems? We saw in our brief discussion of the dynamic-systems models of syntactic change proposed by Niyogi and Berwick (1995; 1997) in §4.2 that variation may spontaneously arise through the interaction of the set of possible grammars, the learning algorithm, and fluctuations in the PLD, but Sapir's point is that the variation over time in a group in relation to the 'linguistic prototype' must be centrifugal. Otherwise, as he says '[o]ught not the individual variations of each locality, even in the absence of intercourse between them, to cancel out to the same accepted speech average?' (160). He proposes drift as the centrifugal force and, arguably, a parametric approach needs something similar.

A final argument in favour of treating diachronic change, seen as parameter change, as something more than a random walk through the range of possibilities defined by UG comes from the observation that the UG-defined space, all other things being equal, is simply too big. For example, if UG contains thirty binary parameters this means that in principle there are 2^{30} , or 1,073,741,824 grammatical systems. Whether this poses learnability problems (a point discussed by Clark and Roberts (1993: 304)) is not the issue here.³¹ Instead, the question is how we can observe synchronic types

³¹ It is not at all difficult to conceive of thirty binary parameters. In Chapter 1, we proposed five main parameters (null subjects, V-to-T, T-to-C, negative concord, wh-movement). The directionality parameter is non-unitary, but breaks up into several sub-parameters related by a markedness convention: I have not stated how many sub-parameters there are (because I don't know), but for the sake of argument we could assume there are ten. Additionally, we have mentioned a parameter governing subject-raising to SpecTP (§3.1.1), four parameters governing auxiliary

or diachronic directions if there are in principle so many systems available. In an earlier paper, Roberts (2001: 90–1), I summarized this point as follows:

the fact that on the basis of a small subset of currently-existing languages we can clearly observe language types, and note diachronic drift from one type to another, is simply astonishing ... languages should appear to vary unpredictably and without assignable limits, even if we have a UG containing just 30 or so parameters. Obviously, we need to find ways to reduce the range of parametric possibilities while retaining (at least) 30 parameters ...

More specifically, the uniformitarian assumption ought to have no force. So many grammatical systems are available that they could appear to vary wildly, and there would be no way to establish uniformity, even with a restrictive UG and thirty binary parameters.

For all these reasons we must add something to the conception that grammars diachronically ‘walk’ around the parameter space defined by UG. Something must cause grammatical systems to ‘clump together’ synchronically in certain areas of that space and to drift towards those areas diachronically. It may be that the force responsible for this is markedness. To phrase things in terms of the theory of dynamical systems, we can think of UG as a state space, a multidimensional space in which each point represents a system-state, i.e. a set of values of parameters, a grammar. This is rather different from seeing the ensemble of the learning algorithm,

selection (§4.1) and a further one concerning the feature-content of *v* (§2.3), a parameter concerning the ability of *C* to Agree for Case with the subject of the TP it introduces (§2.4), a parameter – of rather unclear nature – determining the option of Preposition-stranding (§3.3) and a parameter concerning whether *wh*-expressions are DPs (§4.1). This brings the total to twenty-four. Adding Subject Side and Serial Verbs from the discussion of Baker’s (2001) parameter hierarchy in §3.5, the total comes to twenty-six. Needless to say, it would not be difficult to motivate a further four: the null-subject parameter may break up into at least three parameters, arguably the two originally identified by Rizzi (1982) (see §1.2.1), and a third determining a ‘partial null-subject’ system in the sense discussed in §4.2.6 above. And, of course, many more than this would be needed just to reach observational adequacy. So there is little doubt that thirty is a conservative estimate of the number of parameters. As that number increases, so does the force of the argument about to be made in the text. In our discussion of Guardiano, Gianollo, and Longobardi (2004) in §4.4.5 below, we will see that they propose thirty-six parameters which affect DP-internal syntax only.

the set of grammars, and the probability distribution of PLD items as a dynamical system in the way we saw Niyogi and Berwick (1995; 1997) and Niyogi (2004) do in §4.2.3: here I am simply looking at the set of grammars. Markedness creates basins of attraction, or sinks, in this space, i.e. points towards which grammars always tend to move. In other words, certain areas of the parameter space attract grammatical systems, by being relatively unmarked. Since there may be different markedness conventions and different types of markedness, there may be – and presumably are – various basins of attraction.³²

I do not want to insist unduly on the dynamical-systems terminology here; for present purposes, the notions of attractor, basin of attraction, and space of variation can be taken as metaphorical. The central point is that something like Sapir's notion of drift is required on both empirical and conceptual grounds, even given a highly restrictive theory of UG and a relatively small number of binary parameters of variation, and that this notion can be understood in terms of markedness: markedness defines the areas in the space towards which grammars tend to drift. This view is also compatible with the Inertia Principle: essentially if no other force acts on a grammatical system (i.e. no external contingency radically alters the PLD), a grammar will continue to drift in a given direction – it is conceivable that inertia does not entail stasis.³³ Furthermore, this idea is compatible with Lightfoot's (1999: 90–1) view that parametric change is chaotic in the sense that it is highly sensitive to very small variations in initial conditions. However, Niyogi (2004) shows that the dynamical system constituted by the learning algorithm, the set of grammars and the distribution of PLD is

³² Possibly of various kinds: Lass (1997: 294) describes the distinction between point-attractors, limit-cycle attractors and gutters. Point-attractors resemble a level 'valley' in the state space (or 'epigenetic landscape' (Waddington 1977)) in which a system will come to rest at a steady state. Limit-cycle attractors are like a bowl, around the side of which the system will oscillate. Gutters are sloping valleys, down which the system will continue to move. All of these states could be seen as different manifestations of inertia, in that a system will move along the path of least resistance in the abstract landscape unless its motion is perturbed by some other force. It is an open question whether these distinctions are useful in diachronic syntax, but arguably one worth exploring, as indeed Lass does (1997: 295–304).

³³ Although if we take this view we can no longer construe Inertia as deriving from perfect convergence by language acquirers, as we did in §2.1. It implies that optional convergence, determined by markedness, will force change in a particular direction – towards an attractor.

not chaotic, although it is non-linear. Again speaking metaphorically, a grammar may only need a very small ‘push’ in order to start to drift in a direction which may take it, given enough time, a long distance in the state space from its starting point. In other words, a long-term typological change may come about through a cascade of smaller changes all tending in the same overall typological direction.

In fact, we suggested a slightly more precise and less metaphorical way of thinking about this possibility in our discussion of diffusion and gradualness in §4.1. Given the lexically encoded nature of parameters, as values of formal features of functional heads, there is the possibility of ‘lexical diffusion’ through the functional system: a series of discrete changes to the formal features of a set of functional categories. Such changes might take place over a long period and give the impression of a single, large, gradual change. We can see this as change, for example in the EPP features of functional categories, diffusing through the system of functional categories and following the markedness convention we proposed in §3.5 (57). Something like this might lie behind the changes in word order in the Ethiopic DP that we looked at in the previous section. There is nothing inevitable about such a sequence of changes: it is what happens if no other force acts on the system through the PLD, i.e. if there are no changes in population, contact with speakers of other systems (which can manifest itself in different ways), and speakers’ perceptions of the ‘social value’ of variant properties. Of course, in the actual history of actual languages, these forces do act. As a result, systems are deflected from their ‘path of drift’.

4.3.4. *Cascading parameter changes in the history of English*

It has often been pointed out that English seems to diverge quite radically from the other West Germanic languages. It used to be thought that this had to do with the influence of Norman French, although more recently the effects of Old Norse have sometimes been regarded as responsible for this divergence (see §5.2.2). A series of changes took place in the history of English between 1100 and 1700, which had the net effect of transforming English from a typologically ‘standard’ West Germanic language into the unusual system of Modern English. Arguably, this is an instance of the formal account of drift I have been suggesting here: a cascade of parametric changes diffused through parts of the functional-category system over a

fairly long period of time. It amounts to what Sapir referred to as ‘the vast accumulation of minute modifications which in time results in the complete remodeling of the language’ (1921: 165). Here I want to look at the subgroup of those changes which affected the ENE verb-movement and auxiliary system. (For more detailed discussion of both the empirical and technical aspects of these and the earlier changes, see Biberauer and Roberts (2005b).)

Let us begin with the loss of V2 in the fifteenth century. We can date this change to approximately 1450 (see van Kemenade (1987: 219ff.); Fischer *et al.* (2000: 133ff.)). It is often said, starting with van Kemenade (1987: 196ff.), that this change came about through ‘decliticization’: the earlier V3 orders, where a clitic pronoun intervened between the initial constituent and the verb in C (see Chapter 1, Box 1.2 on V3 in OE; this system persisted in southern dialects of ME, as Kroch and Taylor (1997) show), were reanalysed as involving a non-clitic pronoun in a lower position and the verb in T. So we had the following reanalysis:

- (28) $[_{CP} XP [_C S-cl [_{C[T} V]]] [_{TP} (Subj) t_T \dots]]$
 $[_{CP} XP] C [_{TP} S-pron [_T V] \dots]$

This reanalysis was presumably favoured by the fact that many V2 orders were in any case subject initial, and such orders were prone to be reanalysed as TPs with V-to-T movement (see Adams (1987a, b); Roberts (1993a) on Old French; Willis (1998) on Middle Welsh). Therefore, a consequence of the loss of V2 was that V-to-T movement became a general feature of finite clauses. Where there was a pronoun, the relevant reanalysis was as in (29) (here the set notation indicates that the order of the copy of V and the copy of XP inside TP is irrelevant):

- (29) $[_C XP [_T V]] [_{TP} \{ \dots (XP) \dots (V) \dots \}] > [_{TP} XP [_T V] \dots (XP) \dots]$

This change eliminated V-movement to C and XP-movement to SpecCP in a range of cases, with the corresponding gain in simplicity through the elimination of EPP features on C.

The next change to take place was the lexicalization of T by the reanalysis of modals and *do* as auxiliaries. As we saw in §2.1, this most probably happened ca. 1525–1550 (Lightfoot (1979); Roberts (1985; 1993a: 310ff.); Warner (1997: 382–3)). Roberts (1993a: 310ff.) argues that a further factor in this change was the loss of the infinitival ending on verbs, which had the consequence that constructions consisting of modals followed by an

infinitive were reanalysed as monoclausal: the absence of the infinitival ending meant that there was no evidence for the lower functional T-*v* system. Denison (1985) and Roberts (1993a) suggest that *do* became an auxiliary at the same time as the modals, in the early sixteenth century. Formerly, in Late ME, *do* had been a raising or causative verb (see Roberts (1993a: 282ff.)). This change, again, was a simplification.

The reanalysis of the modals favoured the loss of V-to-T movement. We saw in our discussion of Kroch (1989) and the Constant Rate Effect in §4.1.5 that, although there was variation throughout the ENE period, as Warner (1997: 382–3) observes (see the discussion of his periodization of ENE in §2.1.4), the period 1575–1600 seems to be the crucial one as far as this change is concerned. The development of auxiliaries, and particularly the free availability of ‘dummy’ *do*, including in positive declaratives, meant that T was frequently lexically filled and that this option was always available.

I suggested in §2.1.4, that the loss of morphological expression of the V-to-T parameter created the strong P-ambiguity needed for a reanalysis of the following kind (repeated from (20) of Chapter 2):

- (30) [TP John [T walk-eth] ... [VP ... (V) ...]] >
 [TP John T ... [VP ... [V walks]]]

Following Kroch (1989) and Warner (1997), this reanalysis led to variation for a period, but the innovative, structurally more economical grammar was favoured.

By now the verb-auxiliary system is rather similar to that of Modern English, with the exception of the absence of *do*-support. *Do* could still be freely inserted in positive declarative clauses, and, conversely, clausal negation could appear without *do* in the absence of any other auxiliary, giving rise to examples with the order *not*-V (since V-to-T has been lost):³⁴

- (31) a. Or if there were, it not belongs to you.
 (1600: Shakespeare *2 Henry IV*, IV, i, 98; Battistella and Lobeck 1991)
 b. Safe on this ground we not fear today to tempt your laughter by our rustic play.
 (1637: Jonson *Sad Shepherd*, Prologue 37; Kroch 1989)

³⁴ This is what lies behind the differential rate of change in the introduction of *do* noted in Kroch (1989), which we mentioned in note 14 above.

The development of NE-style *do*-support was preceded by the development of contracted negation, which took place around 1600, as the following remark by Jespersen (1909–49, V: 429) suggests:

The contracted forms seem to have come into use in speech, though not yet in writing, about the year 1600. In a few instances (extremely few) they may be inferred from the metre in Sh[akespeare], though the full form is written.

Around 1600, then, negation contracted onto T, but since V-to-T movement of main verbs had been lost, only auxiliaries were able to be negative. This gave rise to a new system of clausal negation in which negative auxiliaries were used as the basic marker of clausal negation. (It is clear from a range of languages, including Uralic, Latin, Afrikaans, OE, and others, that negative auxiliaries are a UG option.) The new class of auxiliaries included negative modals like *won't*, *can't*, and *shan't*, but also the non-modal negator and *don't/doesn't/didn't*.³⁵ I take these to be the sociolinguistically neutral expressions of negation in contemporary spoken English; Aux + *not* forms with clausal-negation interpretation are ‘expanded’ forms of the negatively inflected aux, which may owe their existence to normative pressures.

Once the negative auxiliaries, including *doesn't*, *don't*, and *didn't*, are established as the unmarked expression of clausal negation (probably by the middle of the seventeenth century; Roberts (1993a: 308)), the modern system of *do*-support comes into being. In this system, merger of *do* in T depends either on the presence of an ‘extra’ feature on T, in addition to Tense and φ -features (i.e. Q or Neg), or on the presence of a discourse effect, in contexts of emphasis and VP-fronting, as in:³⁶

- (32) a. John DOES (so/too) smoke.
 b. He threatened to smoke Gauloises and [smoke Gauloises] he did. ...

With this final development, the present-day English system is in place.

³⁵ Zwicky and Pullum (1983) argue that the negative auxiliaries must be distinct items in the lexicon. The negative *n't* must be treated as an inflectional suffix, rather than a clitic, since inflections but not clitics trigger stem allomorphy, and *n't* triggers such allomorphy (see also Spencer (1991: 381ff.)). The same is argued by E. Williams (1994: 168) on the basis of the unpredictable relative scope relations between various modals and negation.

³⁶ This is connected to Chomsky's (2001: 34) proposal that ‘optional operations can apply only if they have an effect on outcome’.

The series of changes just described can be seen as a cascade of parametric changes. We can summarize them, giving approximate dates for each one, as follows:

- (33) a. loss of V2 (1450) >;
 b. development of lexical T (modals and *do*) (1525) >;
 c. loss of V-to-T (1575) >;
 d. contraction of negation (1600) >;
 e. development of negative auxiliaries (1630s) >;
 f. development of *do*-support (later seventeenth century).

In more technical terms, the parametric changes involved are arguably the following:

- (34) a. (Matrix) C loses EPP-feature triggering V-movement.
 b. Modal and aspectual features of T realized by Merge.
 c. T loses EPP-feature triggering V-movement.
 d. possibly not a syntactic change
 e. Negative features of clause realized by Merge in T.
 f. Obligatory merger of *do* in T restricted to contexts where T has a Neg, Q feature (or clear discourse effect).

All the syntactic changes except (34a) involve changes in the feature make-up of elements merged in T, and there is a clear formal parallel between (34a) and (34c). So we observe exactly the kind of small, incremental changes discussed above, although this is not lexical diffusion, but a series of changes mostly affecting the features of a single functional head. Taken together, they give rise to a major reorganization of the English verb-placement and auxiliary system, and have created a system which is quite unlike anything found elsewhere in Germanic (or Romance), from a starting point in 1400 which was comparable to what we find in Modern Icelandic (see §1.3.1.2, Table 1.1).

As we have seen, English word order changed from OV to VO and from VAux to AuxV in the ME period. In terms of the ‘massive movement’ proposal, VAux order involved movement of vP to SpecTP. The loss of this order is clearly another change in the movement-triggering properties of T. This change must have been a precondition for the loss of V2, since the latter change entailed the development of a V-to-T movement system, which was incompatible with the OV/VAux grammar, if the massive-movement analysis is correct. It is interesting to note that Icelandic underwent the word-order change but not the subsequent changes discussed in the text (see Hróarsdóttir (1999; 2000); Rögnvaldsson (1996); and §1.6.1). It may be that Icelandic never lost V2 because it never had subject (pro)clitics, if the

account of the loss of V2 sketched above based on decliticization is right. Bringing the word-order change into the picture, we see ‘the vast accumulation of minute modifications which in time results in the complete remodeling of the language’ Sapir described. So this is a clear example of parametric drift. This account of cascading parameter change in the history of English is developed in more detail in Biberauer and Roberts (2005b).

What causes the cascade effect? The key idea, due to Lightfoot (1979: 123), is that ‘grammars practice therapy, not prophylaxis’. Essentially, each parameter change skews the PLD in such a way that the next is favoured. We have seen in the description above how each successive change was favoured. In general, then, we see that it is possible to maintain a version of the Inertia Principle and at the same time account for an intricate series of related syntactic changes, not all of which have a purely syntax-external cause. There are many details, both technical and empirical, that remain to be clarified in the account of these changes in ME and ENE, but in general terms we can see this as an example of parametric drift, as described above.

Two final points remain to be made, both stemming from Sapir’s original discussion of drift, and linking up with questions we have looked at elsewhere in this book. First, Sapir remarks (1921: 165) that ‘if this drift of language is not merely the familiar set of individual variations seen in vertical perspective, that is historically, instead of horizontally, that is in daily experience, what is it?’. Here I take Sapir to be asserting essentially that change spreads through orderly heterogeneity in the speech community, in the sense discussed by WLH and in §4.2.2 above. It is clear from Kroch (1989) that there was variation in sixteenth- and seventeenth-century English regarding V-to-T movement and *do*-insertion, and we can assume that the same must have been true regarding the status of the modals and the nature of contracted negation and negative auxiliaries. Presumably, the variants had social value (although it is hard to find evidence of this; they certainly do in most varieties of contemporary English, though – see also Jespersen (1909–49, V: 437)). So in principle we can, in line with Sapir and WLH, tie sociolinguistic variation in with our account of the changes in the ENE verb/auxiliary system discussed above.

A final quotation from Sapir illustrates the second point:

The general drift of language has its depths. At the surface the current is relatively fast. In certain features dialects drift apart rapidly. By that very fact these features betray themselves as less fundamental to the genius of the language than the more slowly modifiable features in which the dialects keep together long after they have

grown to be mutually alien forms of speech ... The momentum of the more fundamental, pre-dialectic, drift is often such that languages long disconnected will pass through the same or strikingly similar phases.

(Sapir 1921: 184)

Here Sapir touches on the question of parallel development; that is, we can observe language families in which several daughters diverge typologically from the parent language, but all in the same way; this is true of the Romance languages in relation to Latin, for example, as has often been pointed out (cf. Harris (1978: 3–17)); in the next section we will observe that Gianollo, Guardiano, and Longobardi (2004) show that the Romance languages are parametrically more similar to each other than any of them are to Latin, at least as regards their DP-syntax. If we think of drift in parametric terms, we can understand this, admittedly rather metaphorical, statement of Sapir's as saying that some parameters are more likely to change than others, given a certain starting point. In the light of Baker's (2001) hierarchy of parameters, discussed in §3.5.4, (60), we could interpret this in terms of the distinction between changes to parameters that are relatively low on the hierarchy and those which are superordinate.

4.3.5. Conclusion

In this section, we have discussed the possibility of parametric drift, considered its relation to typological drift of the more familiar kind, and Lightfoot's (1979) well-known objections to the latter, and sketched a possible example of this drift from the history of English. If parametric change admits of directionality, as I have suggested (*pace* Lightfoot) that it must, the next question to investigate is whether we can exploit this potential directionality in reconstructing lost parametric systems. This is the topic of the next section.

4.4. Reconstruction³⁷

4.4.1. Introduction

In this section, I want to take up a further question from traditional historical linguistics in relation to the parametrically-based approach to

³⁷ This section is based on a talk given with Nigel Vincent at the University of Konstanz in February 1999. I am grateful to Nigel for his collaboration and for extremely stimulating discussions of these issues with him, which greatly clarified my thinking on these questions. The views expressed here are my own responsibility, however.

syntax and syntactic change that I have been presenting here: the question of syntactic reconstruction. Comparative reconstruction is a powerful and effective methodology in historical linguistics; Campbell (1998: 108), for example, says that the ‘comparative method is central to historical linguistics, the most important of the various methods and techniques we use to recover linguistic history’. Its importance for establishing the nature of the phonological systems of unattested languages, or stages of languages, is all but undoubted (Fox (1995) is a useful introduction to the general topic). However, in the area of syntax, the situation is a little different. To quote Campbell again, ‘[o]pinions are sharply divided concerning whether syntax is reconstructible by the comparative method’ (1998: 242). In particular, Lightfoot (1979: 154–66; 1999: 255–7; 2002: 114–30) has argued that the reanalytical nature of parametric change makes reconstruction impossible. On the other hand, Harris and Campbell (1995, Chapter 12) suggest, against the background of a rather different model of syntax but nonetheless an approach to syntactic change which assumes reanalysis as a central mechanism, that syntactic reconstruction is possible. In earlier work (Roberts 1998), I have suggested that parameters can actually be used as a basis for reconstruction. Here I will reconsider the issues involved, and take up the last point in a little more detail. The conclusion will be that syntactic reconstruction using the model assumed here is possible within certain clear limits. Being possible, it is definitely desirable, and sets a clear agenda for future research.

4.4.2. *Traditional comparative reconstruction*

Before entering into the discussion of the issues surrounding syntactic reconstruction, let us briefly review the reconstructive method, using an example from phonology. Campbell (1998: 111ff.) discusses this at some length (see also Fox (1995, Chapter 4)). Consider the following Romance forms, the words for ‘goat’ in various languages:

- (35) Italian: capra /kapra/
 Spanish: cabra /kabra/
 Portuguese: cabra /kabra/
 French: chèvre /ʃɛvr(ə)/

Of course, in this case, we know the original Latin form was *capra* (/kapra/). So we can use the reconstruction of Latin on the basis of Romance forms as

a check that the method is reliable. The method proceeds as follows: once we have a set of cognate forms, we establish the likely sound correspondences. In the case of the forms in (35), we clearly want to say that Italian, Spanish, and Portuguese /k/ and /a/ correspond to French /ʃ/ and /ɛ/ respectively, for example, and that Italian /p/, Spanish/Portuguese /b/, and French /v/ correspond.

The next step is hypothesize what the likely proto-form was, i.e. the form in the ancestor which gave rise to the attested forms in the daughter languages. This is done on the basis of ‘majority rule’ (not because of any predilection for democracy on the part of historical linguists, but rather as a matter of methodological parsimony, since this option involves postulating fewer changes), the likely direction of changes, the factoring of common features, the plausibility of the proposed changes, and the typological plausibility of the overall reconstructed system. Another factor which may be relevant is the relative age of the daughter languages; the older a daughter language is, the more likely it is to be closer to the ancestor (all other things being equal), and so forms in an older language may carry more weight in determining a reconstructed form than those in a relatively younger language. (This criterion does not play any role in the case illustrated by (35), however.)

In the case of the initial consonant, the majority clearly favours /k/, and since /k/ is likely to become palatalized before a front vowel like /ɛ/, we postulate the change from /k/ to /ʃ/ in the development from Latin to French. Majority rule similarly favours /a/ over /ɛ/ as the vowel, here and in the case of the final vowel, for which we posit that /a/ changed to /ɛ/, then to /ə/, and then dropped in final position (Campbell 1998: 121). The factoring of common features is relevant for the medial consonant /p ~ b ~ b ~ v/; the common feature here is labiality, and the natural direction of consonant lenition suggests that the original consonant was /p/, which was voiced in Ibero-Romance and lenited to /v/ in French. Since the /r/ is common to all the cognate forms, we arrive at the reconstructed form /kapra/, which, as I said above, we know to be correct from the Latin record. This brief and sketchy illustration of reconstruction roughly shows how the method works. Much more data is needed to really indicate its effectiveness, and to illustrate the points regarding typology and overall plausibility. But this is enough to allow us to proceed to the questions raised by syntactic reconstruction.

4.4.3. *Questions about syntactic reconstruction*

The first question that needs to be addressed concerns the ontological status of reconstructed forms. When we posit a reconstructed form, what are we actually positing? This question applies to all kinds of reconstruction, of course. There are, roughly speaking, two views on this question as far as phonological and morphological reconstruction are concerned. First, there is what we can call the ‘realist’ view, represented notably by Lass (1997: 270–2). According to this view, accurately reconstructed forms reflect the reality of an earlier stage of a language: what was actually spoken by real people at a given historical moment. Lass (1997: 271) says, of the reconstructed Indo-European /p/ as the initial consonant in the word for ‘pig’, ‘if we met a Proto-Indo-European speaker we’d expect that his word for ‘pig’ would begin with some kind of voiceless labial stop’. The realist view of reconstructed forms can be opposed to the conventionalist view, as favoured by Meillet (1937) and by Lightfoot (1979: 154–66; 1999: 255–7; 2002: 114–30). According to this view, reconstructions simply summarize existing knowledge about relationships among attested stages of languages; there is no commitment as to what the actual historical forms were. In Lightfoot’s words, reconstruction ‘is the exploitation of acquired knowledge to express genetic relations’ (1979: 166).

I will adopt the definition of a proto-grammar, a reconstructed grammar, put forward by Mark Hale (1996: 162). Hale proposes that a proto-grammar is a ‘set of grammars which are non-distinct in their recoverable features’. I take this to mean that, while we cannot claim to be able to isolate a unique proto-grammar, we can approximate it by isolating its recoverable features on the basis of comparison of attested descendant grammars. The recovered features are real features, not simply conventions summarizing existing knowledge about descendant grammars. The features in question are structural features of grammars which are in principle open to variation (or there would be no question as to what to reconstruct); in the domain of syntax, this definition implies the following question: is there any reason to think that parameter values should not in principle be recoverable, in the same way as other grammatical features?

Lightfoot’s argument has always been that reconstruction of syntax is impossible because syntactic change is largely or totally driven by reanalysis and reanalysis is a process with no inherent directionality. In other words, there is no way to say, for example, that an SOV system is more

likely to turn into an SVO one or a VSO one. Because of this, one cannot infer from the present existence of, say, two related languages, one SVO and VSO, what the basic word order of the parent language may have been. As Lightfoot puts it (1999: 257): ‘the kind of reanalyses that occur in catastrophic change constitute cutoff points to reconstruction’. Proto-languages are no more amenable to reconstruction than proto-weather: ‘one can no more reconstruct the syntax of a proto-language than one can reconstruct last week’s weather, and for the same reason: both reflect chaotic systems’ (Lightfoot (2002: 135)).

On the other hand, Harris and Campbell (1995: 353) claim that syntactic reconstruction may be possible provided we can solve the correspondence problem. The correspondence problem relates to the very first step in the method of reconstruction as outlined in the previous subsection: setting up the group of putative cognates whose common ancestor is to be reconstructed. In phonology, this problem is straightforward: yesterday’s segments correspond in some fairly systematic way to today’s (thus French *chèvre* is the inherited reflex (or continuation) of Latin *capra*). But in syntax, we cannot take sentences as the unit of correspondence: of which Latin sentence is *L’état, c’est moi* the continuation, or reflex? As Watkins (1976: 312) puts it: ‘[t]he first law of comparative grammar is that you’ve got to know what to compare’. Harris and Campbell (1995: 347–53) propose a way of solving this problem, and illustrate it on the basis of interesting data from Kartvelian languages. I will not dwell on their approach, however, since it is based on a different set of assumptions about the nature of syntax from those I am adopting here. Their point, however, is well-taken: the correspondence problem appears much more difficult in the case of syntactic reconstruction than in the case of phonological or morphological reconstruction. This is because the phonological and morphological reconstruction are based on a finite array of lexical items, while syntax deals with the unbounded range of sentences. There is little doubt that the sentence cannot be the unit of reconstruction; here the correspondence problem really appears to be intractable.

A third point of view is represented by Watkins (1976: 306). He makes the simple but telling observation that syntactic reconstruction, whatever its methodological basis, has in fact been put to the test, and has passed it very successfully: ‘the confirmation by Hittite of virtually every assertion about Indo-European word order patterns made by Berthold Delbrück [see Delbrück (1893–1900) – IGR] [is] ... as dramatic as the surfacing of the

laryngeals in that language' (Watkins 1976: 306). So one might be tempted not to worry unduly about methodological matters.

Vincent and Roberts (1999) point out two further problems for syntactic reconstruction, in addition to the correspondence problem. One of these is the directionality problem: choosing which attested forms to reconstruct depends on knowing which of the putative changes is the most natural/likely. Do we have a way of deciding which kind of word-order change is most natural, for example? Lightfoot claims that we cannot determine directionality. But we saw in the previous section that it might be possible, and even desirable, to develop a theory of parametric drift, determined by a robust theory of markedness of parameter values. Such a theory may be readily formalizable as dynamical system, and may be compatible both with the Inertia Principle and with Niyogi's (2004) observation that change is non-linear.

The third problem Vincent and Roberts identify is the 'pool of variants' problem. This can be simply illustrated with the forms of the future in some Romance varieties:

(36)	French:	chanter-ai	('I will sing')
	Italian:	canter -ò	"
	Spanish:	cantar -é	"
	Rumanian:	voi cânta	"
	Sardinian:	appo a cantare	"
	Calabrese, Salentino:	no form	

How are we to decide what the original form might have been on this basis? In this connection it is worth pointing out that the Classical Latin future forms *amabo* 'I will love', *dicam* 'I will say', etc., cannot be reconstructed at all on the basis of comparative Romance evidence.³⁸

³⁸ Lass (1997: 273, n. 44) points out that a number of other properties of Classical Latin could not be reconstructed just on the basis of the surviving Romance languages: the 'case-system, verb morphology, OV syntax, three genders, etc.'. This is perhaps slightly overstated: Spanish shows some residue of a neuter gender, and Old French has a residue of the case system. The consistent proclitic placement of unstressed complement pronouns in nearly all Modern Romance languages could conceivably provide a basis for a suspicion of an earlier OV order, although we would need a more predictive theory of syntax than we currently have to state this with any confidence.

Let us now consider the first and third of these issues one at a time. I will leave the directionality problem aside, since it was discussed in the previous section.

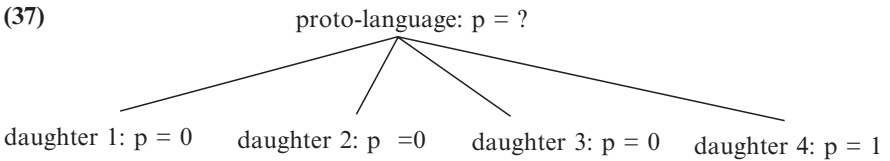
4.4.4. *The correspondence problem*

Let us entertain the idea that identifying the parameters of syntactic variation offers a way to solve the correspondence problem. To make this idea clearer, recall that our guiding assumption throughout this book is that UG contains a finite set of binary parameters which defines the possible variation in grammatical systems. Now for a notational convention: being binary, parameter values can be expressed as members of the set $\{1, 0\}$.³⁹ So, for example, if we state the null-subject parameter as ‘every tensed clause requires an overt subject’ (see §1.2.1) then we can say that languages like French and English have the value 1, while Italian, Spanish, Greek, etc., have the value 0. Here we are doing nothing other than identifying a point of correspondence between the grammatical systems. As we saw in Chapter 1, what can be and has been achieved in the synchronic domain using parameters can be applied to the diachronic domain. In accordance with our demonstration in §1.2.2, that the null-subject parameter changed its value between Old French and Modern French, then, we can say that Old French had the value 0 for the statement of the parameter just given, while Modern French has the value 1. Again, we have identified a point of correspondence between Old and Modern French. There is no conceptual difficulty here at all.⁴⁰

³⁹ Following a notational device standard in formal semantics, we can then think of parameter values as the truth values of contingent statements about grammars.

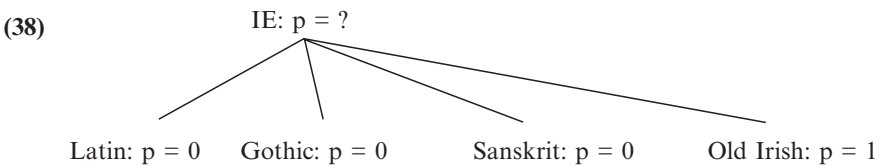
⁴⁰ Lightfoot (2002: 128–30) argues that what is suggested in the text (which was first suggested in Roberts (1998)) does not help with syntactic reconstruction, since we need an approach to directionality. He then goes on to mention four parameters which he considers cannot stand in a markedness relation. Three of these, OV vs. VO, V-to-T movement, and the presence or not of English-type auxiliaries, have been extensively discussed here and a view as to which is the marked value has been suggested (see §3.5.1). The fourth parameter concerns the bounding of wh-movement, as put forward by Rizzi (1982, Chapter 1). However, the status of this parameter is at present somewhat uncertain – see Rizzi (1989).

To illustrate how we might proceed with ‘parametric reconstruction’, consider the case where we have four daughter languages with determinate values for a given parameter, and we wish to reconstruct the value of that parameter in the parent language. The situation is schematized in (37):



Here we would be justified in postulating $p = 0$ in the parent language on two grounds. First, ‘majority rule’ (or more precisely, analytical economy): this hypothesis entails that only one of the four daughters changed, and is therefore the simplest. This is justified as long as we know that the daughters with $p = 0$ are older than the one with $p = 1$. Also, if we know that $0 > 1$ is the more natural change for p (the directionality question) than $1 > 0$, postulating $p = 0$ is justified. Considerations of typological naturalness might also play a role. In other words, exactly the considerations that were at work in our illustration of phonological reconstruction above are relevant here; there are no differences of principle in syntactic reconstruction.

To put the discussion on a slightly more concrete basis, suppose that p is the VO/OV parameter, stated for simplicity and concreteness as ‘V{precedes(1)/follows(0)} the direct object’.⁴¹ And suppose that we are trying to reconstruct the value of this parameter in Indo-European on the basis of what we know about Latin, Gothic, Sanskrit, and Old Irish. Then we have the situation in (38):



⁴¹ Of course, this is unlikely to be the correct formulation of this parameter in practice – see §2.5.

Majority rule clearly favours $p = 0$ (the OV setting) here. So do several other considerations: (i) Latin, Gothic, and Sanskrit are all older than Old Irish;⁴² (ii) Archaic Old Irish shows some evidence of OV orders (see Doherty (2000a, b); Russell (1995)); (iii) OV > VO seems to be a more natural change than VO > OV. On the basis of the criteria normally used in comparative reconstruction, then, we would come to the conclusion that $p = 0$ was the Indo-European value of this parameter. This agrees with the discussion in Brugmann and Delbrück (1897–1916); W. Lehmann (1993: 187ff.), and the references given in the latter; and Fortson (2004: 142).

Pursuing this idea, we can also define the notion of a ‘continuation’ of a form. We can say that a continuation is a potentially variant property which is stable across grammars in the diachronic domain. Applying this to parameters, this is the case where a parameter value is ‘correctly’ set by acquirers over a given period. In other words, continuation is the opposite of abductive reanalysis (and represents the straightforward case of Inertia, leaving aside the possibility of ‘inertial drift’). In terms of the illustration just given, we would say that $p = 0$ in Latin, Gothic, and Sanskrit is the continuation of the Indo-European value of this parameter. In other words, over many generations (two or three millennia, in fact) the relevant trigger experience was sufficiently robust for acquirers to have continued to correctly set this parameter value on the basis of their PLD.

The above considerations suffice to establish that parameters can be used as a solution to the correspondence problem for syntax. If parameters are formal features of lexical entries, this is no surprise. Lexical items can be used as corresponding items in general: those relevant for parameters are simply somewhat more abstract than the non-functional lexicon. Given this, and our general remarks on directionality, I conclude that there is no problem of principle in using parameter values as features of a proto-grammar in Hale’s sense as defined above. We can thus now state the following:

- (39) A proto-grammar is a set of grammars which are non-distinct in their recoverable parameter values.

⁴² Most of the Old Irish corpus dates from the eighth and ninth centuries (Mac Eoin, 1994: 102) Gothic from about four centuries earlier.

We can illustrate (39) by reconsidering some of the parameters we have been looking at in the preceding chapters. (40) are partial restatements of (54A, C, E, F) of §3.5.1):⁴³

- (40) a. T licenses a null subject. Italian: 1; English: 0
 b. (Root) C attracts V. German: 1; English: 0
 c. C attracts WH. English: 1; Chinese: 0
 d. v attracts VP. German: 1; English: 0

Now, if we impose an arbitrary order on the sequence of parameter values (for example, that given in (40)), we can give the index of each grammatical system, as follows:

- (41) English: 0010
 Italian: 1010
 Chinese: 1001
 German: 0111

The correspondence problem can now be restated in a maximally general form as: for some proto-language P, what is P's index regarding the properties listed in (40) or some comparable list of parameters? For illustration, let us briefly consider what these might have been for Indo-European.

First, the null-subject parameter. There is very little doubt that Indo-European was a null-subject language. All the oldest attested daughters are null-subject languages, and so we can simply assume that nothing changed in this regard. To the extent that a positive value for the null-subject parameter is connected to 'rich' verbal inflection, the traditional idea that IE was a language with a rich verbal inflection (see Szemerényi (1996: 233); Fortson (2004: 83)) supports this conclusion. We have already seen reasons to treat IE as OV (see (38)). So we can immediately state (42), where the asterisk indicates that the value of the parameter in question is uncertain:

- (42) Indo-European: 1**1

Concerning V2, there is some evidence for an optional verb-fronting rule in Indo-European, and clear evidence – possibly the best known fact about

⁴³ On Chinese, see Huang (1984) for null subjects, Huang (1982) for wh-movement, Huang (1989) for V-movement, and Li (1989); Sybesma (1999) for detailed discussion and analysis of Chinese word order. The values of the other parameters given here in English, Italian, and German were discussed in §1.1, §1.3, and §2.5.

Indo-European syntax – that various ‘light’ elements, pronouns, particles, and light adverbs, were attracted to the second position (Wackernagel 1892; see also Fortson (2004: 146–7)). If we identify this second position with (root) C, then we observe that Indo-European may have had a kind of category-neutral version of V2. This is easily expressible in the formal system assumed here: the EPP feature of C which attracts heads is simply not specified for V. So it is possible that Indo-European had the positive value for a parameter closely related to the V2 parameter.

Finally, it is widely agreed that Indo-European had *wh*-movement in both interrogatives and relatives to a position in the left periphery of the clause (see Hale (1995) on Sanskrit; Garrett (1990) on Hittite; Kiparsky (1995) on Archaic Latin and various older forms of Germanic; and Fortson (2004: 145)). If this is correct, then we can conclude, with the proviso that the ‘second-position’ parameter did not relate specifically to verbs but to heads more generally, that the ‘parametric index’ of Indo-European for the four parameters listed in (40) was (43):

(43) 1101

Although many questions of data and analysis remain open, this conclusion shows that reconstruction of some of the parameters of Indo-European is at least possible. In this way, a potentially interesting research agenda emerges, since all the issues raised above need to be clarified across the range of older Indo-European languages.

4.4.5. *The ‘pool of variants’ problem*

Let us now return to the ‘pool of variants’ problem. We saw in (36) that the forms for the expression of the future in various Romance languages and dialects show quite a range of formal options, and it is at first sight rather unclear which of these forms may represent the Latin original. However, our approach to markedness can point us in the right direction, even if it does not provide an exact conclusion.

Let us assume that the representation of future concerns the realization of a tense or mood feature, associated either with T or *v*. The French, Italian, and Spanish forms in (36) all express this feature synthetically: they realize it by means of movement of the inflected verb into the functional system. Rumanian and Sardinian, on the other hand, have periphrastic

constructions in which it is plausible to think that the auxiliary is merged in the relevant functional position. Given our approach to markedness, these represent the less marked option. The Southern Italian dialects in fact represent a still less marked option, in that there is no formal expression of the future feature at all. Since these are the most conservative dialects, they may tell us more than the others. From this we can conclude that the original Latin future form, if it had one, has been lost; and of course we know that this is true. Here our approach is useful in a rather negative way, but it is possible to say a little more. Since the Merge option is less marked than the Move one, we might conclude that the French, Italian, and Spanish forms derived from an option that was originally more similar to the Sardinian and Rumanian one. Given the evidence that Rumanian has undergone extensive influence from the neighbouring non-Romance Balkan languages, all of which have a future auxiliary diachronically derived from the verb ‘to want’, we can give preference to the Sardinian variant as the reconstructed form. Hence we could reconstruct an original periphrastic future formed with a ‘have’ auxiliary. All the future forms, whether synthetic or analytic, clearly preserve the verbal infinitive, and so we can think that the common ancestor forms consisted of a ‘have’ auxiliary combined with the infinitive. This is the correct reconstruction for Vulgar Latin (Harris (1978: 145–6); Roberts and Roussou (2003: 48–58) and the references given there).⁴⁴ As we mentioned above, we could never reconstruct the Classical Latin *-bo* or *-am* forms, but we can see how our markedness-based approach can help with the pool-of-variants problem.

4.4.6. *Parametric comparison*

In a very interesting recent series of papers, Longobardi (2003), Guardiano and Longobardi (2003; 2005) and Gianollo, Guardiano, and Longobardi (2004) (henceforth GGL) develop a method of measuring grammatical differences between languages: *parametric comparison*. Their starting point is the observation that, while languages differ from one another in

⁴⁴ This is not to imply that Sardinian preserves intact the Vulgar Latin form. A future of almost exactly this form is found at all attested stages of Sardinian, but, intriguingly, lacking the *a* element between *aere* (‘have’) and the infinitive at the earlier stages, which suggests it might be a syntactically rather different entity.

all aspects of their structure, some pairs of languages differ from each other more than others do: Spanish and Portuguese are very similar to each other indeed; English is quite similar to German and more similar to Dutch, while Japanese is significantly unlike all of these languages. Many, if not all, of these degrees of structural and lexical difference can be correlated to historical relationships. But no systematic way of quantifying our intuitions about grammatical relatedness has emerged. GGL's work attempts to fill this gap. Its relevance to historical syntax concerns the fact that if, as GGL claim, the method is more reliable than attempts to quantify lexical, morphological, and phonological distance (on these, see in particular the papers in MacMahon (2005); MacMahon and MacMahon (2003); Nakleh, Ringe, and Warnow (2005); Nakleh *et al.* (2005); Ringe, Warnow, and Taylor (2002), and the brief summary of the last of these below), then it may be possible to use parametric comparison to establish or confirm relationships among languages.

GGL restrict themselves to a specific syntactic domain, arguing that this controls the number of variable properties and makes the cross-linguistic data more comparable and manageable. The domain they choose is the DP. In §1.6.1 we observed that within the DP there is variation in Article-Noun order, Plural marker-Noun order, Noun-Relative Clause order and Noun-Genitive order (see (129j–m) of that chapter). Among the other properties that may vary are the presence/absence of number marking (English: YES; Japanese: NO), the presence/absence of a system of articles (English: YES; Japanese: NO), and the presence/absence of a system of classifiers (English: NO; Japanese: YES). GGL extend and systematize these observations, producing a *parameter grid* illustrating the values of thirty-six parameters concerning the internal structure of DP across eighteen languages, shown in Table 4.2.

This grid gives an enormous amount of information regarding the DP-internal syntax of the languages in question, far more than can be summarized in the space available here (GGL give more details). The important point for present purposes, however, concerns the range of similarities and differences in parameter values that we find in grids like this. The central point is summed up nicely by the following quotation:

1/2 has no probative value at all. However, ... it is easy to increase the number of *comparanda*: 30 binary independent parameters generate 2^{30} languages = 1,073,741,824. Now the probability for two languages to coincide in the values of 30 chosen parameters = 1 in 2^{30} , of three languages 1 in $(2^{30})^2$, i.e. less than one in a

Table 4.2 Parameter grid for nominal syntax

	It	Sp	Fr	Ptg	Lat	CIG	NTC	Grk	Got	OE	E	D	Big	SC	Rus	Heb	Ar	W
1. ± gramm. interpr. gender	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
2. ± gramm. gender -1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0+
3. ± gramm. number -2	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+	0+
4. ± number. on N (BNs)+1 or +2 or +3	+	+	-e	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
5. ± ambiguous singulars +4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
6. ± gramm. count ('null article')	+	+	+	+	-	-	-	+	-	-	+	+	-	-	-	-	-	-
7. ± gramm. def in DP	+	+	+	+	-	+	+	+	+	+	+	+	-	-	-	+	+	+
8. ± gramm. non-anaph. def. +7	+	+	+	+	0	+	+	-	+	+	+	+	-0	-0	+	+	+	+
9. ± strong ref/def in D +7, +8	+	+	+?	0	+	+	+	0	-?e	-	-	+	+	0	0	+	+	-?
10. ± enclitic def/delctic	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-
11. ± art + PP +art	-	+	-	+	0	+	+	-	0+	0+	-	0+	0	0	0	-	-	-
12. ± gramm. prox. in D +7	-	-	-	-	0	-	-	-	-	-	-	-	0	0	-	-	-	+
13. ± high dem.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+
14. ± low dem.	-	+	-	-?	+	+	+	+	+	-	-	-?	-?	-?	+	-?	-	-
15. ± def. check. dem.+7,(+13 or (-6, +14)	+	+	+	+	0-	-	-	+	+	+	+	+	0-	0-	+	-	+	+
16. ± hend initial DP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
17. ± def. checking Aps	-	-	-	-	-	-	-	+	-	-	-	+	+	-	-	-	-	-
18. ± def. spread on Aps +7 or +17	-	-	-	-	0-	-	-	-	-	-	-	-	-	0-	+	+	-	-
19. ± def. spread on Mod +7 or +17	-	-	-	-	0-	+	+	-	-	-	-	+	-	0-	-	-	-	-
20. ± ordered As	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
21. ± inversely ordered As +20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	0
22. ± free APs in Mod +20	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-?	-?	0+	0+
23. ± context. infl. A +infl. As	-	-	-	-	-	-	-	+	+	0-	+	+	+	+	-	-	-	-
24. ± N over ext. arg. -20 or -21	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0+	0+	+
25. ± N over GenO +24	+	+	+	+	-	-	+	+	+	0-	+	+	+	+	+	0+	0+	+
26. ± N over low As +20, +25	+	+	+	+	0-	0-	-	-?	-	0-	-	-	-	-	-	0+	0+	0
27. ± N over M2 As +26	+	+	+	+	0-	0-	0-	0-	0-?	0-	0-	0-	0-	0-	0-	0+	0+	0
28. ± N over high As +24/28	-	-	-	-	0-	0-	0-	0-	0-	0-	0-	0-	0-	0-	0-	0+	0+	0
29. ± Consistency Princ. +24/28	+	+	+	+	?	?	?	+	?	+	+	+	-	-	-	0	0	0
30. ± prepos. Gen (vs.infl.Gen)	+	+	+	+	-	-	-	-	-	+	+	+	-	-	-	+	+	+?
31. ± free Gen (non-agr.) -30	0+	0+	0+	0+	+	+	-	-	-	0+	0+	0+	-	-	0+	0+	0+	0+
32. ± GenO +30 or -31	-	-	-	-	0+	0+	+	+	+	+	+	+	+	+	-	-	-	+
33. ± GenS +30 or (-31, +32)	-e	-e	-	-e	0+	0+	-	-	+	+	+	+	-	-	-	+	+	-
34. ± possessive pronouns	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	-	-	+
35. ± oblig. possessive pronouns +34	+	+	+	+	+	-	-	0-	?	?	-	-	+?	+?	+	0-	0-	+
36. ± def checking poss. +7, +34	-	+	+	-	0-	-	-	0-	0-	+	+	+	-	0-	0-	0-	0-	-

Source: Gianollo, Guardiano, and Longobardi (2004: 7).

billion of billions. But even choosing just 10 such parameters, the probability that their values coincide in three languages is less than one in a million.

(Guardiano and Longobardi 2005: 155)

Hence, if we can observe significant coincidences in parameter values across languages, then we are very likely to be observing some non-accidental relationship. The greater the number of parameters that coincide

in value and the greater the number of languages in which we observe the coincidence, the more likely it is that we are observing a genuine relationship. GGL argue that it is overwhelmingly likely to be the case that the relationships so observed are due to common historical ancestry rather than to contact, although the latter possibility cannot be completely excluded. In other words, parametric similarities reflect retention of parameter settings across generations, i.e. successive convergence over many generations by acquirers. (This can be attributed to the Inertia Principle.)

Evolutionary biologists have developed computational methods for constructing phylogenetic relationships on the basis of data regarding shared characteristics. The programs that construct such relationships are, naturally, indifferent to whether the data they are organizing concerns shared linguistic traits or shared genetic features (both can in fact be regarded as polymorphisms). Details regarding some of these programs are given in MacMahon and MacMahon (2003: 29–32). Such programs can present the phylogenies they construct as trees, similar to those familiar in historical linguistics. The data in Table 4.2 was implemented by GGL using the program *Kitsch*, one of the *PHYLIP* programs developed by Felsenstein (2001). GGL's implementation of their data using *Kitsch* gave the results shown in Figure 4.5.

The tree is rooted by convention, and so all the languages appear to be related. What is more relevant is what we find on the lower branches. Here we see that the Semitic, Germanic, Romance, and Slavonic families are each correctly grouped together to the exclusion of other languages. Also, the varieties of Greek form a subgroup. Here, though, we see that Latin is in this group, which is incorrect. (This may be connected to the question of parallel development mentioned at the end of §4.4.4 above.) But, by and large, the groupings replicate the known historical relationships among the languages in question. Of course, these groupings are well-known as the historical relations among the languages considered, and so nothing new is being discovered here. However, the fact that the parametric-comparison method can independently reveal the correct historical relationships provides an initial indication that it is a valid method for detecting historical relationships.

GGL further argue that an approach to measuring relatedness which relies on parametric syntax has certain advantages over an approach based on lexical similarities. The first of these is the discreteness of parameter values: the values of a parameter do not form a continuum or cline of any

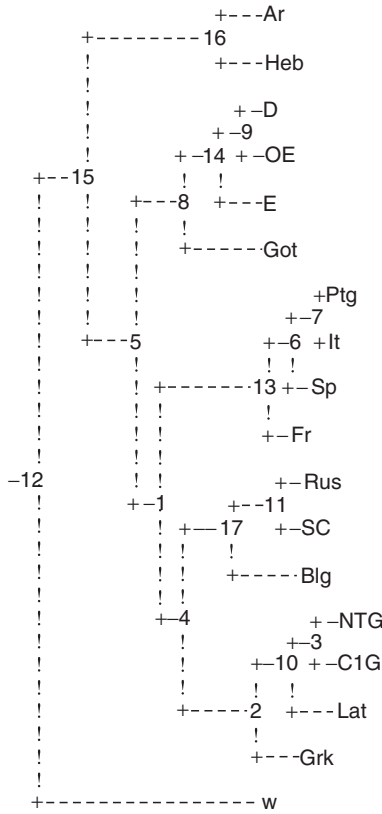


Figure 4.5 Result of application of Kitsch to the data in Table 4.2

Source: Gianollo, Guardiano, and Longobardi (2004: 14).

kind. On the other hand, the meanings of individual lexical items are often difficult to delimit. Second, the binary nature of parametric oppositions means that we have the maximally simple range of possibilities of variation in any given case. Third, the number of parameters is finite, and in fact rather small, usually thought to be more than twenty but less than 100; hence the total range of comparanda is of a manageable size. Finally, there is in principle no uncertainty of comparanda: we are always sure when we are comparing like with like in the domain of syntax. Again, in the domain of lexical comparison, this is frequently a difficulty.

In this connection, it is worth comparing GGL’s approach with that in Ringe *et al.* (2002). These authors used techniques from evolutionary biology to try to identify the first-order subgrouping of Indo-European languages. A central concept in this connection is that of a character, which they define as ‘an identifiable point of grammar or lexical meaning which

evolves formally over the course of the language family's development, ... each state of the character ought to represent an identifiable unique historical stage of development – a true homology [shared trait inherited from a common ancestor – IGR]' (71). An example of a lexical character would be English *hand* (=1), Ger *Hand* (=1), Fr *main* (=2), It *mano* (=2), Rus *ruká* (=3); the numbers are arbitrary values designating each state of the character. An example of a phonological character would be the sequence of changes including Grimm's Law, Verner's Law, initial-syllable stress, merger of unstressed **e* with **i* except before **r*. If these changes are absent, the character has value 1, if they are present, it has value 2. (Unsurprisingly, this character singles out Germanic among the subgroups of Indo-European.)

Ringe *et al.* used a database of twenty-four languages representing all ten Indo-European subgroups, and compared 322 characters (twenty-two of them phonological, fifteen morphological, and the rest lexical). The result of running the tree-optimization software was that eighteen characters were incompatible with the best tree. As they put it: 'in computational terms our result is a total failure' (86). However, since fourteen of the eighteen incompatibilities involved Germanic, they simply removed the Germanic languages from the comparison. This gave a more satisfactory result, and one according to which Italo-Celtic, Balto-Slavonic, the *satem* group, and Graeco-Armenian emerge as IE subgroups. Nakleh, Ringe, and Warnow (2005) developed a system of phylogenetic networks, rather than trees, in an attempt to reconstruct contact relationships in addition to genetic inheritance relationships. In this system, they were able to include Germanic, postulating ancient contact with Balto-Slavonic and Italo-Celtic.

Given GGL's observations regarding the advantage of parametric comparison over lexical, phonological, or morphological comparison, a natural proposal would be to combine the two approaches, and treat parameter values as characters, thereby adding syntax to the structural comparison. In fact, parametric grids like that in Table 1.2 (§1.3.2.2) can be taken to indicate the syntactic characters. In addition to the DP-related parameters looked at by GGL, one could readily consider several of the clause-level parameters we have been looking at throughout this book as characters (for example, second-position phenomena, wh-movement, null subjects, the presence of clausal particles such as question particles, OV vs. VO order and the associated properties, the presence or absence of pronominal

clitics, auxiliaries, etc.); most of these are relevant to the older Indo-European languages. In this way, syntactic properties could play a real role in reconstructing the relationships among the Indo-European languages.

However, Ringe *et al.* observe that there are two things which must be avoided if the phylogenetic method they adopt is to work. The first of these is what they call ‘backmutation’, the case where a change takes place in one direction, followed, at some later stage in the development of the family, by its exact reversal. In the areas of the lexicon, morphology, and phonology, they observe, backmutation is ‘either improbable or vanishly rare’ (70). They say that ‘we simply do not find cases in which the contrast between two elements A and B in a structured system is eliminated from the language, then . . . reintroduced in precisely the same distribution that it originally exhibited’ (70). This seems to be clearly true of such changes as phonemic split and merger, the loss or gain of inflection, and changes in word-meaning. However, we have seen that it is probably not true of syntactic parameters. (Recall the discussion of the null-subject parameter in the history of French and various Northern Italian dialects in §1.2.1). However, we may also observe that the parametric system as a whole does not undergo backmutation. To avoid this problem, it may be necessary to consider sets of parameters as characters (just as Ringe *et al.* propose treating sets of sound changes as phonological characters).

The second thing which must be eliminated in order for phylogenies to be computed, according to Ringe *et al.*, is parallel development (in genetic terms, this is analogy rather than homology). As we have seen, the development of the Romance languages may be a case of parallel development. One way to avoid this problem might again be to take sets of changes rather than individual changes as evidence for groups of related languages (or clades). In any case, it seems doubtful that syntactic change poses any problems not already encountered, perhaps in a slightly different form, in the area of phonology.

A further obvious difficulty in using syntactic parameter values as characters is that much less is known about the syntax of a number of older IE languages compared to their phonology, lexicon, and morphology. However, here there is no issue of principle (beyond the usual difficulties of working with extinct languages): we simply need to apply the analytic, descriptive techniques of syntactic theory to the analysis of the older

languages, as has to some extent already been done by Garrett (1990), Hale (1995), and Kiparsky (1995).

In general, then, parametric comparison can give us a way to quantify grammatical differences. As such, it may play a major role in developing our theories of language typology and language change. It may also shed light on major questions in historical linguistics. In this connection, Longobardi (2003: 2) observes that we can pose ourselves five fundamental questions in linguistic theory, as follows:

- (44) a. What are the recorded samples of human linguistic behaviour?
 b. What are the actual human languages?
 c. What are the biologically possible human languages?
 d. Why do we have precisely these actual languages?
 e. Why do we have precisely these biologically possible human languages?

The answers to these questions correspond to different levels of adequacy in linguistic theory. The first three correspond to **observational, descriptive, and explanatory adequacy** as defined in Chomsky (1964). The fourth, Longobardi calls ‘actual historical adequacy’, while the fifth relates directly to the goals of the Minimalist Program as described in Chomsky (2002; 2004; 2005a, b). Concerning historical adequacy, Longobardi points out that ‘one of the best examples of historical explanation ... in the human sciences has been provided by the historical-comparative paradigm in linguistics, where reconstructed protolanguages play the role of initial conditions deriving observed languages in conjunction with some general hypotheses about possible linguistic changes’ (2003: 4). He further urges that we should ‘take advantage of the combined insights of the two major scientific revolutions in linguistics, those which gave rise respectively to the historical-comparative paradigm during the XIX century and the “synchronic-cognitive” paradigm in the XX. It is such a combination that may yield substance to a good deal of the historical-explanatory program’ (2003: 5). It is in this way that the parametric approach may take its place in the long tradition of historical and comparative linguistics.

4.4.7. Conclusion

In this section I have considered the possibility of using a parametric system for syntactic reconstruction. It is possible to deal with the three problems identified by Vincent and Roberts (1999) in terms of the general

approach to change we have been advocating throughout. We can also define a notion of proto-grammar in parametric terms, following Hale (1996), and provide tentative but fairly detailed reconstructions of certain specific parameters of Indo-European. There is no reason not to take a realist view of the partial reconstructions offered here. I conclude from this that syntactic reconstruction is possible, even in a principle-and-parameters reanalysis-driven approach to change (*pace* Lightfoot (1979; 1999; 2002)). Moreover, we have seen how the parametric-comparison method developed by GGL may offer the possibility of confirming and establishing historical relationships among languages.

4.5. Conclusion to Chapter 4

As I said at the beginning of this chapter, the purpose of the discussion here has been to consider to what extent the general model for syntactic change introduced in Chapter 3 poses problems for, is compatible with, or is even useful or insightful for a number of questions concerning what can be rather generally thought of as the dynamics of syntactic change. These questions mainly concern ‘external’ aspects of change, and how these appear to be incompatible with the kind of ‘internalist’ account of change that was developed in the preceding chapters.

What has emerged is that the parametric approach does not really pose any special problems for reconciling the external aspects of change with what I take to be the necessarily internalist account of how systems change. Concerning the question of reconciling the apparently gradual nature of change with the fact that parametric change must be seen as instantaneous, we observed that several factors may ‘cushion’ the effects of a parametric change: sociolinguistic factors relating to diffusion, microparametric change and variation, and the possibility of competing grammars and/or true formal optionality. In §4.2, I looked in some detail at WLH’s discussion of sociolinguistic factors in change, and how notions like diglossia, social stratification, and ‘ordered heterogeneity’ might be compatible with a parametric approach. We saw how integrating these aspects into a parametric approach would in principle not be problematic, and how Kroch’s notion of grammar competition may or may not be helpful here. In §4.3, I argued, *pace* Lightfoot (1979; 1999: 208–10) that ‘parametric drift’ probably does exist and should in any case be allowed by our theory. A clearer

understanding of markedness should provide an account of this phenomenon. Finally, in §4.4, I argued that syntactic reconstruction using parameter values as the point of correspondence is possible and may be desirable; again, this is contrary to what was argued in Lightfoot (1979; 1999; 2002).

In general, a parametric approach to language change does not really raise any special problems, or make any special predictions, as far as issues of the type considered here are concerned. (Again, this conclusion differs from what was argued in Lightfoot (1991; 1999), where it is claimed that the adoption of a parametric model makes specific predictions regarding the speed of change and other matters.) The main contribution that a parametric approach has to offer to historical work does not concern the nature of change, since, as Hale (1998) and Longobardi (2001) both point out in different ways, parametric models are inherently atemporal and change is of course a temporal notion. Instead, what a parametric approach has to offer is a firm analytical foundation for the analysis and comparison of grammatical systems; it is a precise, rigorous, and flexible formal system for representing the knowledge and acquisition of syntax in the individual, and for representing the dimensions along which that knowledge can vary across individuals or aggregates of individuals. It provides an essential basis for the analysis of syntactic change, while at the same time not prejudging any of the more difficult, external questions which have been raised in this chapter.

In the final chapter, I turn to the most complex ‘external’ question of all, that of how language contact should be thought of in parametric terms, and to what extent it is responsible for observed syntactic changes.

Further reading

Historical and Indo-European linguistics

Campbell (1998) is a very good general introduction to historical linguistics, with the emphasis on phonological and morphological change. **MacMahon (1994)** is a very useful general introduction to historical linguistics, with a good balance of discussion of syntactic, morphological, and phonological change, as well as some intriguing discussion of the historical connections between the study of language change and the study of evolutionary change in organisms/species. **Durie and Ross (1996)** is a collection of articles on the

nature and status of the methodology of reconstruction and its usefulness in historical linguistics. **Lightfoot (2002)** is a rebuttal of the claims regarding syntactic reconstruction made in Harris and Campbell (1995) and, to a lesser extent, **Roberts (1998)**. The latter is a review of Harris and Campbell (1995) (see the further reading to the Introduction), in which an early version of the idea discussed in §4.4.4 that parameters can solve the correspondence problem for reconstruction is put forward. **Vincent and Roberts (1999)** argues, in agreement with Harris and Campbell (1995) but on different grounds, in favour of syntactic reconstruction. The discussion of the ‘pool-of-variants’ issue in §4.4.4 relies heavily on this paper. **Gianollo, Guardiano, and Longobardi (2004)** is the most detailed discussion and illustration to date of the method of parametric comparison. An impressive amount of data is collated and summarized in their survey of thirty-six parameters in eighteen languages. They conclude that this method may be more useful than lexically-based quantitative methods as a tool for discovering long-distance relations across languages. Earlier papers exploring the same idea are **Longobardi (2003)** and **Guardiano and Longobardi (2003; 2005)**. **MacMahon (2005)**; **MacMahon and MacMahon (2003)**; **Nakleh, Ringe, and Warnow (2005)**; **Nakleh et al. (2005)**; and **Ringe, Warnow, and Taylor (2002)** pursue the question of quantifying relatedness using lexical and, in the case of the latter three articles, phonological and morphological information. Nakleh, Ringe, and Warnow (2005) develops Ringe, Warnow, and Taylor (2002) in considering phylogenetic networks rather than simply trees, in an attempt to reconstruct contact relations. These two papers arrive at a very interesting reconstruction of the first-order relations among the Indo-European languages. **Hale (1998)** is a general article on the generative approach to diachronic syntax. A number of very useful conceptual clarifications are made, notably regarding the atemporal nature of grammars. **Benveniste (1966)** is a classic study of the expression of possession in relation to perfectivity. Although the focus is on Classical Armenian, a number of important typological observations are made regarding the cross-linguistic incidence of ‘have’-like verbs and auxiliaries. **Denison (2003)** is a very useful critical summary of the explanatory role of the S-curve in historical linguistics. **Ellegård (1953)** is a classic study of the incidence of *do* in fifteenth-, sixteenth- and seventeenth-century English, whose results are used extensively by Kroch (1989), and rather less extensively by Roberts (1985; 1993a). **Horrocks (1997)** is a wide-ranging and useful account of the history of all aspects of the structure of Greek

from the earliest times up to Modern Greek. **Jespersen (1909–49)** is a monumental study of the historical syntax of English, containing invaluable data and discussions of an extremely wide range of phenomena. **Roberts and Kato (1993)** is a collection of articles dealing with the diachronic syntax of Brazilian Portuguese, many of which focus on the question of the apparently changing status of null subjects in contemporary Brazilian Portuguese. **Penny (1991)** is arguably the main history of Spanish written in English in recent years. **Pintzuk (2003; 2005)** continues the line of research in Pintzuk (1991; 1999) (see the further reading to Chapter 1) in looking at word-order variation and change in the history of English in terms of grammars in competition. These two articles look in detail at different types of OV constructions in OE and ME. **Biberauer and Roberts (2005b)** is a recent extension of the earlier work on word-order change by these authors (see the further reading to Chapter 2), which relates the ME word-order changes to the changes in Late ME and ENE discussed in §4.3.3. **Santorini (1989; 1992; 1993)** deals with word-order change in the history of Yiddish, arguing for a competing-grammars approach to these phenomena. **Fontana (1993)** is a detailed study of word order in Old Spanish, in which it is argued that this language, like a number of other Romance languages, went through a V2 stage. **Lakoff (1972)** looks at the development of the expression of modality and sentential complementation, concentrating primarily on Latin. It is a rare example of work on syntactic change in the generative-semantics framework. **Mathieu and Sitaridou (2005)** looks at the ‘hyperbaton’ construction of Classical Greek, arguing that it is a case of general left-branch extraction. **Brugmann and Delbrück (1897–1916)** is a reprint of the classic summary of the state of the art in Indo-European linguistics at the end of the nineteenth century. **Wackernagel (1892)** is a classic study of one aspect of the syntax of the Indo-European languages, the central observation of which is that in many of the older Indo-European languages various ‘light’ elements (pronouns, light adverbs, and sentential particles) show a strong tendency to appear in the second position in the clause. Wackernagel’s observations are still highly relevant to the study of the ‘left periphery’ of the clause, perhaps to be construed as CP, and remain largely unexplained in satisfactory theoretical terms. **Watkins (1976)** is a very interesting discussion of the possibility of reconstructing Indo-European syntax, featuring a very intriguing proposal regarding relative clauses. **Hale (1995)** looks at word order in Sanskrit in government-and-binding terms, arguing that this language has a fixed clause structure in which topics

precede foci, both of which precede the ‘core’ of the clause (TP in the terminology used here). This analysis has been influential and has been applied to a number of other older Indo-European languages – see Fortson (2004: 141–6) for discussion. **Szemerényi (1996)** is another very thorough introduction to Indo-European linguistics. **Fortson (2004)** is a recent and very useful textbook on Indo-European. Somewhat unusually for works of this kind, a whole chapter is devoted to syntax.

Syntax

Chomsky (1993) (also published as Chapter Three of Chomsky (1995)) is the earliest published statement of the Minimalist Program, and contains many of the central ideas of this program. The technical details, however, are rather different from the more recent versions of the Minimalist Program. **Borer (1984)** is an early study of clitics and the construct-state possessive construction in Modern Hebrew, notable for the first proposal that parameters should be stated as parts of lexical entries. **Freeze (1992)** argues that ‘have’-type verbs and auxiliaries, expressing possession, obligation, existence, location, and the perfect, are systematically and universally related to ‘be’-auxiliaries. This idea was taken up by Kayne in his (1993) study of auxiliary-selection in Romance. (This is reprinted in Kayne (2000); see the further reading to Chapter 2.) **Mahajan (1994)** proposes an account of the observation that ‘have’-auxiliaries tend not to be found in OV languages, which connects with a general account of the ergative parameter (the choice between ergative-absolutive and nominative-accusative case/agreement marking – see Box 4.1). A number of extremely interesting similarities between auxiliary-selection and ergative case/agreement marking are pointed out. **Müller (2004b)** proposes a highly original analysis of ‘ergative parameter’, the distinction between ergative-absolutive and nominative-accusative languages, invoking the ordering of operations in the highly derivational model of Chomsky (2001). **Li (1989)** is a detailed study of Case-assignment and word order in Mandarin Chinese, using the technical assumptions of government-and-binding theory. **Sybesma (1999)** is a more recent study of word order and clause structure in Mandarin Chinese, which develops and proposes alternatives to a number of the ideas in Li (1989). **E. Williams (1994)** is an essay on morphosyntax, in which, among other things, a very lexicalist approach to morphology is argued for.

Morphology and phonology

Aronoff (1976) is an important work on morphology in generative grammar, which is in many respects still relevant to current theoretical concerns. **Spencer (1991)** is a very thorough study of pre-optimality-theory morphological theory. **Zwicky and Pullum (1983)** put forward an influential account of the differences between cliticization and affixation. Some of the ideas put forward here are taken up in Spencer (1991: 381–4). **Halle (1962)** is one of the first articles to look at language change from a generative perspective, proposing an approach to sound change in terms of an early conception of generative phonology. Abductive reanalysis through language acquisition plays a central role in this approach. **Kiparsky (2003)** is a very interesting and relatively recent discussion of phonological change. The discussion of markedness reversal in relation to lexical diffusion is extremely useful, and almost certainly relevant for our understanding of syntactic change.

Italian dialects

Cocchi (1995) is a treatment of auxiliary-selection in Central and Southern Italian dialects, in terms of an early version of minimalism. **Loporcaro (1998)** is a very detailed study of past-participle agreement in Romance, from both a synchronic and a diachronic point of view, using the theoretical framework of relational grammar. **Manzini and Savoia (2005)** is a monumental study of the phonology, morphology, and syntax of Italian and Rhaeto-Romance dialects. **Tuttle (1986)** is a diachronic study of the person-driven auxiliary systems of the Central-Southern Italian dialects, while **Vincent (1982)** looks at the general development of the perfect periphrasis from Latin to Romance.

Sociolinguistics

Weinreich, Labov, and Herzog (1968) is the foundational text for the sociolinguistically-oriented study of language change, particularly sound change in progress. Many of the ideas first put forward here have been developed in greater detail in Labov (1994; 2001). The article is most

relevant and useful in demonstrating how any general theory of the structure of language must take ‘external’ factors into account if a full picture of how historical changes take place is to be obtained. **Labov (1966)** is a seminal study in sociolinguistics, looking at the incidence of postvocalic /r/ in the English of New York City, in which it is shown that this element functions as a sociolinguistic variable, in that its incidence correlates with the age, gender, and social class, of speakers. Furthermore, the variation across age-groups, classes and genders indicates that this variety of English is moving from lacking postvocalic /r/ to having it. Finally, the incidence of postvocalic /r/ in this variety shows it to be a prestige form for many groups of speakers. **Labov (1972)** is a collection of important articles, several of them on Black English Vernacular (the English of the Afro-American population of the United States, now usually referred to as African-American Vernacular English (or AAVE)). These include a seminal article on negative concord in this variety, which is briefly summarized in §4.2.2. **Labov (1994; 2001)** are the first two volumes of a major work on the nature of language change, whose emphasis is primarily on sound change, and which brings together many of Labov’s concerns going back to his earliest work in the 1960s. At the time of writing, a projected third volume is about to appear. **Muysken (2000)** is a detailed study of code-switching (or code-mixing), with emphasis on the widespread existence of subsentential code-switching.

Other important works

De Saussure (1959) is a classic of twentieth-century linguistics. More than any other, this work was responsible for shifting the emphasis of linguistic theory away from historical linguistics and towards synchronic study. De Saussure was also responsible for the concept of the linguistic sign as an arbitrary relation between a signified and a signifier, the idea of the structure of language as a system of oppositions, and an early version of the concept of the phoneme. **Sapir (1921)** is another classic of twentieth-century linguistics and a founding text of American structuralism. In it, the idea that all languages are equally complex and equally worth studying is argued for in great detail on the basis of data from a range of Amerindian languages. Sapir also proposes an elaboration of Schleicher’s (1866) morphological typology.

5

Contact, creoles, and change

Introduction	383	5.3. Creoles and creolization	406
5.1. Second-language acquisition, interlanguage, and syntactic change	384	5.4. Language creation in Nicaragua	427
5.2. Contact and substrata	389	5.5. Conclusion to Chapter 5	438
		Further reading	440

Introduction

In the previous chapters I have argued for the existence of parametric change in the historical record (Chapter 1), for its general utility in accounting for types of syntactic change (Chapter 2), and for an acquisition-driven model of parameter-setting and change (Chapter 3). In Chapter 4, the discussion centred on how this approach to syntactic change may or may not shed light on a number of issues connected to the dynamics of change, concluding that the parametric approach has little that is special to contribute to these questions *per se*, but that the interest of the approach is that it provides a formally rigorous yet flexible analytical tool for comparative syntax. The purpose of this final chapter is somewhat different: the goal here is to discuss second-language learning and acquisition as factors in syntactic change and, more generally, issues related to the phenomenon of language contact, and how this may affect the PLD in a way that causes parametric change. We briefly discussed contact in relation to parametric

change in §3.3.1; however, the discussion there was only concerned with illustrating how the PLD could change in such a way that the Regress Problem could be solved. In particular, we did not discuss the nature and role of ‘imperfect learning’ of a second language by adult learners; this and related issues, including the nature of substrate effects and the question of the status of creoles are the main issues we will look at in this chapter.

§5.1 looks at interlanguage and second-language acquisition. In §5.2 I consider contact and **substratum** phenomena. §5.3 considers **creoles** and **creolization**. Here we consider two approaches to creoles: one is based on the idea that Bickerton’s (1981; 1984) Language Bioprogram causes all parameters to be set to unmarked values, while the other is based on the idea that creoles show radical **relexification** (usually by a European language) of a substrate system which retains syntactic features of the original language or languages of the creole-speaking group. The interest of creoles lies in the fact that children whose PLD consists largely of **pidgin** seem to be faced with a particularly extreme instance of an impoverished stimulus for acquisition, and so how they cope with this may be of great interest for our general understanding of how children handle underdetermined input, which I take to be a feature of all first-language acquisition. Against this background, I will also consider DeGraff’s recent arguments against what he calls ‘creole exceptionalism’ (see DeGraff (2003; 2004; 2005; to appear)). In §5.4 we look at a genuinely ‘exceptional’ situation, and what might be the most spectacular example of impoverished PLD ever documented. This is the case of a group of deaf children in Nicaragua who apparently invented their own **sign language** over a period of a few years, documented by Kegl, Senghas, and Coppola (1999) and further discussed in Kegl (to appear). The structural properties and the ‘history’, insofar as it can be traced, of this language are clearly of the highest theoretical interest. We will look at these questions, and consider what conclusions may be drawn for linguistic theory.

5.1. Second-language acquisition, interlanguage, and syntactic change

The topic of the acquisition of a non-native second (or third, etc.) language has given rise to a fairly significant research literature in recent years: the references given in R. Hawkins (2001) and White (2003) attest to this. In

this section, I do not propose to provide an introduction to this field; R. Hawkins and White do this. Instead, I want to present an overview of the issues which are relevant to understanding the phenomena of language contact and substrate phenomena in the diachronic domain, which are the subject matter of the next subsection. To some extent, issues related to second-language acquisition (henceforth L2A) are also relevant to creoles and creolization. More generally, L2A is relevant to syntactic change since, as we have mentioned, imperfect learning of a second language by adults in a contact situation may affect the PLD of a subsequent generation and lead to contact effects of various kinds. We noted in Chapter 3 that this is one way in which the Regress Problem can be solved. So our goal here is to look at what the concept of ‘imperfect learning’ really means.

In §3.1 we saw how first-language (L1) acquisition is conceptualized in terms of principles-and-parameters theory. There we saw that first-language acquisition is the process by which the language faculty passes from the state S_0 , corresponding to UG, to the stable state S_S of adult competence. Schematically, L1A falls into several stages $\langle S_0, \dots, S_{i>0}, \dots, S_{n>i}, S_{n+1<S}, \dots, S_S \rangle$. We suggested that the intermediate stages correspond to an immature competence, in which the process of fixing the parameters of the grammar to an approximation of the grammar underlying the PLD is ongoing.

L2A differs from L1A in a number of respects. In fact, we can observe differences in all three aspects of the acquisition process: the initial state, the intermediate states, and the final state. The initial state differs from that of L1A by definition: the L2 acquirer does not start from UG, but already has acquired the L1 grammar. One might therefore expect that this grammar influences the L2 grammar and the process of L2 acquisition in various ways. The extent to which this is true is one of the central points which is debated in the research literature on L2A. Whatever the correct point of view on this matter, we can see that there is a clear difference with L1 acquisition.

The intermediate states of L2 acquisition differ in principle from those of L1 in up to three main ways. First, they may be affected by the L1 grammar, just as the initial state is. Second, the course of L2 acquisition may not be the same as that of L1 acquisition. Third, since it seems to be a matter of casual observation that L2 acquisition is very frequently imperfect, it may be that the acquisition process ‘fossilizes’ at an intermediate state; an intermediate state may thus amount to a final state. To this one

can also add that the time course of L2 acquisition may vary much more than that of L1. Whilst it is usually said that L1 acquisition begins in the first year of life and is complete by the age of six or so, L2 acquisition can be a much more protracted affair, with even highly accomplished L2 speakers only approximating true L1 competence after many years of practice in and exposure to the L2. (See the discussion of Coppieters (1987) and Birdsong (1992) in White (2003: 252–4).)

Finally, the final state of L2 acquisition very often fails to correspond to the target. Although one of the central ideas of this book is that this may also be true of L1 acquisition, and this is what underlies much syntactic change, it is clear that L1 acquisition usually approximates the target very closely, so much so that much work on L1 acquisition assumes perfect convergence, and the Inertia Principle can be seen as reflecting this. But the outcome of L2 acquisition is very often conspicuously divergent in relation to the target. In fact, explaining this is one of the central issues in L2A research.

Behind all of these differences between L1A and L2A lies the question of the critical period, the idea that ‘there is a time period which is optimal for language acquisition, with a maturational decline with increasing age’ (White 2003: 245); we will see some very interesting evidence for the critical period in §5.4). L2A takes place after the putative critical period, and it is plausible to think that at least some of the differences in the intermediate states and in the final state are due to this. In fact, if we imagine that fixing UG parameters involves effectively fossilizing them, then once L1 acquisition is complete no further language acquisition of any kind can take place and we would expect to find the differences between L1A and L2A just described. We will see directly that interlanguage studies have provided evidence against such a strong view of the critical period.

If it is correct to think of the critical period in terms of the atrophying either of UG, perhaps for the reason just given, or of the parameter-setting device (the learning algorithm), then it follows that L2A can have no access to UG. In that case, we expect L2 production and comprehension to mirror UG only via the L1 competence, and we might expect L2 production to be ‘wild’, in the sense that it might include structures which are not allowed by UG. Furthermore, we do not expect to find evidence of parameter-setting in L2A of the kind that we expect in L1A. Studies of L2A and interlanguage competence – the production and comprehension capacities of L2 learners and speakers – have centred on this question, as documented by both R. Hawkins (2001) and White (2003).

Again, we can divide the acquisition process into the three types of state in order to see more clearly what is involved. If UG is not involved in L2A, then the initial state of L2A must be either the L1 grammar or a non-UG-determined cognitive state, perhaps due to ‘general learning’ strategies of some kind. The former point of view characterizes the ‘local impairment’ approach of Beck (1998); the latter the ‘global impairment’ approach of Clahsen and Hong (1995). On the other hand, if UG is involved in L2A then the initial state can again be the L1 grammar, or it can be UG itself, i.e. equivalent to the initial state of L1A. The former point of view is that of the Full Transfer Full Access approach (Schwartz and Sprouse (1996)) and the No Parameter Resetting approach (for example, Smith and Tsimpli (1995)). The latter point of view is known as Full Access (without Transfer), and has been advocated notably by Flynn and Martohardjono (1994) and Flynn (1996).

The intermediate states can be characterized by parameter-resetting, on a UG-based approach: this is the view of both the Full Access, Full Transfer approach and the Full Access (no Transfer) approach. Alternatively, one can think that parameters cannot be reset – this is the view taken by Smith and Tsimpli. If UG plays no role in L2A then the possibility of parameter-resetting does not arise, and general-learning strategies must be at work; this of course does not rule out the possibility that an able L2 learner might be able to ‘simulate’ fluent, native-like linguistic behaviour in performance, but would by definition not have true competence in the L2.

Regarding the final state, the UG-based theories all agree that an L2-like grammar can be attained, while the non-UG-based theories claim that this is impossible. As mentioned above, we would then expect L2 performance to reflect UG properties only via the L1, if at all, and we would expect ‘wild’ structural features to emerge in L2 production, at least.

White (2003, Chapter 4) considers the evidence for the various points of view just given from experimental studies of various kinds regarding the nature of L2 interlanguage performance. She concludes:

On the whole, the results are consistent with Full Transfer Full Access: learners start out with L1 functional categories, features, and feature strength and are able to acquire L2 categories, features, and feature strength. Nevertheless, some aspects of the results are puzzling for this view, in that effects of the L1 reveal themselves more often than not in the form of variability, with L1 and L2 properties co-occurring, rather than there being initial effects of the L1 setting alone.

(White 2003: 148)

(‘Feature strength’ here refers to the movement-triggering property in the version of minimalism in Chomsky (1993; 1995): it is extensionally equivalent to the EPP features of the current theory as assumed here). If parameters represent underspecified features which simply have to be set to a constant value by the learning algorithm, then we might think that L1 and L2 differ in that the consistency requirement does not obtain in L2. It may be that the learning algorithm does not really do its job in the L2A case, but only once, for L1. Atrophy of the learning algorithm may also be one way to construe the critical-period hypothesis.

Two other aspects of L2 performance are relevant for our purposes here. First, it is widely observed that L2 performance is better in syntax than in morphology, particularly inflectional morphology. White comments that it ‘is well known that L2 learners exhibit optionality or variability in their use of verbal and nominal inflection and associated lexical items’ (2003: 178). If morphology is a cue for parameter values (see §3.3.3), then we can see how L2 performance may affect the PLD for a subsequent generation. Second, although L2 acquisition frequently fails to match the target grammar, there is no reason to assume that it does not correspond to a possible instantiation of UG, i.e. a grammar in the sense of having the parameters set (although possibly not to determinate values). To quote White again: ‘steady-state interlanguage grammars are not wild’ (2003: 266).

It is quite probable that the question of the UG-based status of interlanguage is not of central importance for understanding contact phenomena in diachrony. It may simply suffice to observe that ‘imperfect learning’, which could be construed in terms of any of the accounts of interlanguage sketched above, is enough to perturb the PLD of subsequent generations in such a way as to bring about syntactic change, and thereby solve the Regress Problem. In principle, different types of contact situation and different patterns of change could help decide among the various approaches to interlanguages, but in practice we have too little crucial data both regarding interlanguage and regarding the contact situations and the changes caused by them to be able to do this in the current state of knowledge. One could also think that non-UG-based interlanguage might be ineffectual in bringing about change, if acquirers ignore non-UG-based input. However, certain views of creolization take pidgins to be non-UG-based systems, and it is clear that the Deaf Nicaraguan children who invented a new language did not have UG-based input, so it

seems that linguistic behaviour which does not have a basis in UG can function as a somewhat impoverished form of PLD.

However, we can see that interlanguage is probably UG-based, and has three important properties: (i) it is variable, in that certain parameters seem inherently unstable; (ii) inflectional morphology tends to be lacking; and (iii) it does not correspond exactly to the target. We can take these three points as the content of the notion of imperfect learning. All of these factors can significantly perturb the PLD. Consider a situation where an older generation (or cohort of speakers) speaks an interlanguage with these properties, owing, perhaps, to a foreign invasion, and so this forms (part of) the PLD for the next generation (or cohort). It is intuitively clear, but also has been shown in the computational models produced by Niyogi and Berwick (1995; 1997) and Niyogi (2004), that the PLD will be such that it is able to trigger an L1 grammar which differs in some parameter values from the L1 of the older generation/cohort, despite the fact that the latter group may have successively converged on its L1 grammar on the basis of the PLD from the previous cohort prior to the contact situation. This, then, is how contact may affect PLD. We see that the conclusions regarding the likely nature of interlanguage are quite important for our understanding of how contact can bring about syntactic change.

Let us now turn to some concrete examples of contact and substratum effects.

5.2. Contact and substrata

5.2.1. Introduction

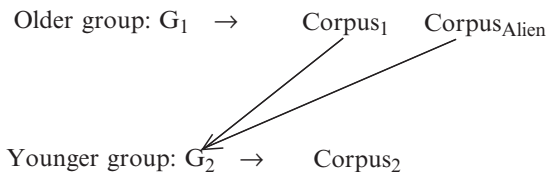
There is no doubt about the existence of the phenomenon of language contact, and its potential importance for understanding language change. Thomason (2003: 687) says ‘most of what historical linguists study under the designation “language change” is due to contact’. Moreover, Kroch (2000: 716) describes contact as an ‘actuating force for syntactic change whose existence cannot be doubted’. We have already discussed the role it may play in changing the PLD, and thereby solving the Regress Problem, in §3.3 and in the previous section. My goal in this section is to look more closely at the role of contact in affecting the PLD, and thereby acting as an

‘actuating force’, to use Kroch’s expression. I begin by reintroducing the distinction between direct and indirect contact, introduced in §3.3, making some modifications in the light of the discussion in §4.2. I will then discuss the possibility that the change in English word order from OV to VO, which, as we have mentioned, may have been caused by contact with speakers of Old Norse in the Danelaw (see in particular Kroch, Taylor, and Ringe (2000); Trips (2002)). I will illustrate the notion of substratum in relation to varieties of English spoken in Wales and Ireland, considering the possibility that certain syntactic peculiarities of these varieties may be attributable to a Celtic substratum (Cottell 2002; Kallen 1994; Thomas 1994). Finally, I will consider the typology of contact relations put forward by Thomason and Kaufman (1988) in the light of the general approach I have been developing here.

I have pointed out that contact may bring about parameter-resetting because where two grammatical systems are in contact, the PLD is affected by an alien grammatical system. What this means is that Generation 2 in the schema for abductive change in (20) of Chapter 4, or more precisely, the younger group of speakers, is subjected to a different kind of PLD from Generation 1, or the older group.¹ The younger group receives PLD that either directly or indirectly reflects a distinct set of parameter values from that which underlay the PLD for the older group.

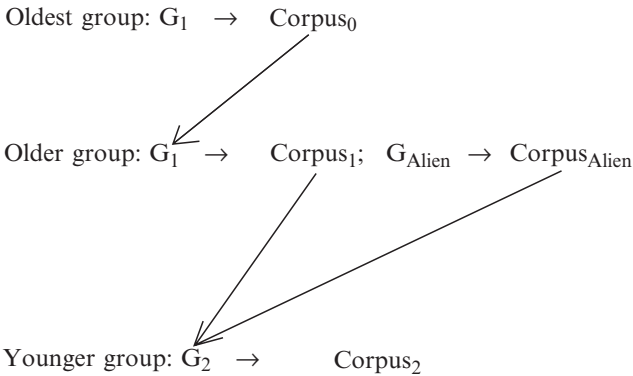
We distinguished the direct and indirect cases of language contact influencing PLD as in (1) (modified from Chapter 3, (23)):

(1) a. Direct contact:



¹ From now on, I will use the terms ‘older group’ and ‘younger group’ rather than ‘Generation 1’ and ‘Generation 2’, given the discussion of Weinreich, Herzog, and Labov’s (1968) critique of the notion of intergenerational change in §4.2.5.

b. Indirect contact:



The core of the distinction between direct and indirect contact, as (1) shows, is that in the direct case the PLD simply contains a significant quantity of tokens from a distinct system, while in the indirect case the older group uses a second language in interaction with the younger group. Here the PLD for the two groups is very obviously distinct. The distinction between direct and indirect contact corresponds approximately to that which is sometimes made between borrowing and imperfect learning, respectively (see for example Kroch (2000: 716)), although the direct case does not really have to result in ‘borrowing’ on the part of either the younger or the older group: the Alien Corpus merely has an effect on the younger group’s PLD.

5.2.2. *Contact and word-order change in the history of English*

I will now take up the possibility that an important change in the history of English may have resulted from contact between Old Norse (ON) and OE in the Danelaw in the ninth to eleventh centuries. This is the change from OV to VO, which has figured several times in our discussions by now.² This idea is pursued in some depth by Trips (2002), who shows that one

² The idea of syntactic influence of ON on OE is touched on by Jespersen (1938: 76), who states that ‘the intimate fusion of the two languages must certainly have influenced syntactical relations’. Regarding word order, he mentions (77) the fact

North-Eastern ME text, the *Ormulum*, written probably in Bourne, Lincolnshire, around 1180 (Burnley (1992:79)), shows a number of syntactic features that are typical of the Scandinavian languages but not typical of OE. Trips concludes that ‘the contact situation between Scandinavian and English brought about a number of syntactic changes, especially the word order change from OV to VO’ (333).

The contact between OE and ON in the Danelaw was probably of an indirect kind: the older group in the schema in (1b) can be thought of as Norse immigrants who imperfectly learned OE. This adult-acquired OE formed part of the PLD for the younger group, who were native OE speakers. The PLD for the younger group was thus clearly different from that of the older group, whose native language was ON, and from that of an older group of OE speakers, or that of OE speakers outside the areas of Danish immigration, settlement, and intermarriage, whose PLD was not affected by ON.

The extent of Scandinavian influence on English has often been commented on, for example, by Jespersen (1938: 60–77) and Trips (2002: 11–2, 65). There was undoubtedly a high degree of lexical borrowing: elements of core vocabulary such as *sky*, *skin*, *die*, *ill*, *ugly*, etc., and, more strikingly, prepositions and functional items such as *from*, *at*, the plural forms of the present tense of *be*, and – perhaps the best known case – the 3pl pronouns *they*, *them*, and *their*, are lexical items of Scandinavian origin which must have come into English through contact with native speakers of ON at or shortly after the period of the Danelaw. (This idea is entertained by Kroch and Taylor (1997: 317–8).) There is also reason to think that the verbal inflections of Northern OE and ME were simplified due to contact with ON. If pronouns can be borrowed, perhaps word-order parameters can be changed; after all, on the assumptions being made here, word-order parameters are specified as formal features of functional categories such as *v*, while pronouns are the realizations of formal features of the functional category *D*. Hence there is little difference, as far as the grammatical system

that ME and Scandinavian show Genitive-Noun order, while OE usually had the opposite, but adds ‘in these delicate matters it is not safe to assert too much, as in fact many similarities may have been independently developed in the two languages’. Also, Mitchell and Robinson (1992: 133) comment that ‘the influence of the dialects spoken by the Danish invaders of the ninth century could have made itself felt and may well have been more advanced in colloquial OE than in the more conservative forms of the language recorded in the manuscripts’.

is concerned, between one and the other. So the hypothesis certainly has an initial plausibility.

Following the hypothesis that contact with speakers of ON was responsible for the actuation of the change from OV to VO in English, we must assume that ON was a VO language (Trips 2002: 331). This idea can be entertained independently of the nature of the parameter(s) responsible for VO order. The older group's imperfect learning of OE meant that they spoke VO OE, or at least produced a larger number of surface strings with this order owing to the nature of their second-language competence in OE than native OE speakers would have done. (Recall the complexities of OE word order, as mentioned in §1.6.2.) This would have sufficed to give acquirers of OE, the younger group in (1b), exposure to PLD favouring a VO grammar for OE. Thus the imperfect learning of OE by Norse immigrants who were native speakers of VO ON sufficed to provide an 'alternate form' of OE to a given group of acquirers of OE in the ninth/tenth-century Danelaw. We thus arrive, through the schema in (1b), at stage (1) of the change as described by Weinreich, Labov, and Herzog and discussed in §4.2.5: 'a speaker learns an alternate form'. In our terms, a speaker acquires a grammar with an innovative parameter value. Stage (2), whereby 'the two forms exist in contact within his competence' may correspond to a period of competing grammars, as documented by Pintzuk (1991; 1999) and described – with the difficulties noted there – in §4.2.3. This stage could also in principle correspond to a single grammar with a new parameter value allowing for the option of surface VO order alongside OV; I will not explore this alternative here, but it is developed in Biberauer and Roberts (2005a). Stage (3) of the change from OV to VO – when the OV system becomes obsolete – is independent of the question of Scandinavian influence and so is not relevant to our present concerns. Stage (1) took place in the ninth and tenth centuries, and is not directly attested. Stage (2) is reflected in the variation we see in the Early ME texts (see Kroch and Taylor (2000) on this) and in the differences between ME dialects in the area of the Danelaw and those spoken elsewhere: Trips' (2002) study of the *Ormulum* supports this conclusion. It is important to see that contact with ON only introduces a new variant; presumably acquirers of OE in the Danelaw were also exposed to PLD produced by native speakers of VO OE which was able to trigger an OV grammar – this is what leads to the variation at stage (2).

This account relies on two assumptions: one concerning the grammatical structure of the ON spoken in the Danelaw, and the other concerning the sociolinguistic situation in the Danelaw and in early ME: the idea that ON speakers imperfectly learned OE and thereby altered the PLD for subsequent generations. The first assumption is that ON (or the variety of it spoken in England) was VO, or at least more predominantly VO than OE. The second assumption is that Norse settlers in the Danelaw learned OE as a second language under historical circumstances which meant that the PLD which resulted from their linguistic behaviour significantly influenced the word-order patterns of the language in later generations. Let us now look at each of these assumptions in turn.

First, ON word order. We have no direct evidence at all of the word order of the ON spoken in England, as this was never written down; see Kastovsky (1992: 331): ‘practically no Scandinavian manuscripts exist’. We must therefore consider what is known about the older attested stages of North Germanic more generally, and draw what conclusions we can from that evidence.

Whilst all the Modern North Germanic languages are VO, Old Icelandic was OV (cf. Hróarsdóttir (1999; 2000)). Rögnvaldsson (1996) showed that all attested stages of Icelandic from the earliest texts (the Family Sagas, written in the thirteenth and fourteenth centuries) up to the early nineteenth century show either pure or mixed OV order; in fact, Rögnvaldsson (1996: 69–72) argues for a competing-grammars analysis for this long period of the history of Icelandic. (Recall that both Hróarsdóttir and Rögnvaldsson use the term ‘Old Icelandic’ to include ON; see Chapter 1, note 39.) Delsing (2000: 271) shows that Old Swedish, up to 1300, was predominantly OV. Faarlund (1994: 64–7; 2004b) says that ON had mixed word order, with OV more common in subordinate clauses and in poetry. He says ‘it may be that Old Scandinavian was still underlying OV’ (66). He also observes that Ancient Scandinavian, the language of the early runic inscriptions up to the seventh century, was OV. Since the ON of the Danelaw represents an older stage than the earliest attested ON, which dates from the thirteenth century, it is possible that it was ‘more OV’ than ON, both in its structural features and in the frequency of OV order. Thus, the assumption that the ON spoken in the Danelaw was entirely or predominantly VO, and so significantly different from OE in this respect that contact with ON was the actuating force in the shift from OV to VO in English generally, does not seem to be supported by what little evidence is

available. The most plausible hypothesis seems to be that that variety of ON was a ‘mixed’ OV system, where ‘mixed’ could mean either that there were competing grammars, or that the system allowed true optionality. In other words, there is no reason to think that ON was significantly different from OE as far as the OV/VO parameter is concerned. Therefore, contact with ON cannot have actuated word-order change in the way described above, and assumed by Trips.

Let us now turn to the second assumption behind the contact-based account of English word-order change: that ON speakers learned OE as a second language under circumstances which brought about their influence on the PLD for later generations. We can also make an argument based on the external historical facts, as far as they are clear, against the idea that contact with ON was the actuating force in word-order change in this way. Examples like (17) from Chapter 4, repeated here for convenience, show, as Pintzuk (2002: 282) says, that late ninth-century West Saxon was already a mixed variety:

- (2) a. He ne **mæg his agene aberan.**
 he NEG can his own support
 ‘He cannot support his own.’
 (CP 52.2)
- b. hu he **his agene unðeawas ongietan wille**
 how he his own faults perceive will
 ‘how he will perceive his own faults’
 (CP 22.21–2)

Recall that (2a) illustrates AuxOV order and (2b) OVAux. As we have seen, Pintzuk interprets this variation, and variation between OV and VO orders at the same period, as evidence for grammars in competition, in other words as part of an ongoing word-order change. But the word-order variation observed here cannot be attributed to ON influence, since the translation of the *Cura Pastoralis* was written in King Alfred’s lifetime (Burnley (1992: 17), i.e. before 899, while the earliest date for Norse settlement in the Danelaw is 865 (Sawyer (1971); in Thomason and Kaufman (1988: 276); Kastovsky (1992: 322); and the references given there). The Norse settlers were certainly fairly few in number to begin with (whatever their later numbers – see below) and only occupied a part of the country distant from Wessex, where the *Cura Pastoralis* was translated. It is highly unlikely that the influence of ON contact, especially if this was due to imperfect learning of English by Scandinavians, could have spread from the North-Eastern

portion of the country where the Scandinavians mostly settled, to Wessex in just thirty-four years, which is the largest window of time the facts permit. So it seems that ON simply did not have enough time to have actuated the grammar competition that led to the word-order change.

One might think that, although both ON and OE were ‘mixed’ OV/VO systems, the contact between the two engendered something akin to a creolization situation. If creoles tend to show unmarked values of parameters, then we could think that this gave rise to the shift to the unmarked VO value of the parameter. (Recall the postulation of unmarked values for parameters in §3.5.1.) However, Thomason and Kaufman (1988: 263ff.) argue very convincingly against this, concluding that ‘[w]hen all the relevant data are examined . . . it is apparent that a creolization hypothesis is not required to explain the facts of Northern Middle English, nor is it even likely’ (265). They also point out that ‘Norse influence could not have modified the basic typology of English because the two were highly similar in the first place’ (266–7). I conclude that the idea that ON influence in the Danelaw functioned as an actuating force in the change from OV to VO in the history of English cannot be supported by the available evidence regarding ON word order, and that the possibility of creolization of OE as a result of contact with ON is not realistic. External facts such as the very short window of time for Norse influence to make itself felt on Alfredian West Saxon OE, especially given the initially small numbers of Norse settlers in the Danelaw and the physical separation of the Danelaw from Wessex, also militate against this idea.

Let us consider the evidence regarding contact between ON and OE more carefully. Thomason and Kaufman (1988: 263ff.) survey the available evidence in detail. There is no doubt that Norse speakers learned English: Thomason and Kaufman state that ‘Norse probably lasted no more than two generations after 955’ (267), and ‘we are convinced that Norse was largely or entirely absorbed by English in the Danelaw by AD. 1100’ (282). Therefore the idea that the PLD for language acquirers may have been affected in the manner described above by the imperfect learning of OE on the part of native speakers of ON has some initial support.

ON speakers settled in England between 865 and 955, probably not in very large numbers.³ The evidence for large-scale immigration is uncertain

³ This and the comments in the rest of the paragraph on the nature of the settlement in the Danelaw are based on the extensive quotations from Sawyer

(see also the comments in Kastovsky (1992: 324)); what is clear is that Viking soldiers and their families settled in various localities over a period of about a century. The densest area of settlement was Leicestershire, Lincolnshire, Nottinghamshire, and Yorkshire (see also Burnley (1992: 416)). It is likely that the Scandinavians mixed with the native population everywhere they settled. The contact between ON and OE must have taken place, then, between roughly 865 and 1100 primarily in the counties mentioned, i.e. in the North-East Midlands and Yorkshire.

Thomason and Kaufman (1988: 282–304) discuss in detail what they refer to as the ‘Norsification’ of OE. They claim to deal with ‘*all* the discoverable grammatical influence of Norse on Old English (apart from syntactic rules)’ (304, emphasis theirs). Since it doesn’t deal with ‘syntactic rules’, their survey cannot provide direct information on the situation regarding word-order change, but it does provide some very useful indications concerning the effect of the contact between ON and OE on the OE and Early ME grammatical systems. They define what they call the ‘Norsified dialects’; these are the varieties spoken in the counties just mentioned as those most densely settled by Scandinavians, along with Northern East Anglia, parts of Northamptonshire and Cambridgeshire, Cheshire, Lancashire, and Derbyshire.⁴ The Norsified dialects ‘have not only heavy lexical influence from Norse, but also have adopted a significant number (between 24 and 57) of Norse derivational and inflectional affixes, inflectional processes, and closed-class grammatical words’ (283). They argue that this variety originated in the North-East Midlands, specifically in the areas identified as those mostly heavily settled by Scandinavians, and that it significantly influenced Northern ME. (In fact, they argue that it had more influence on Northern ME than did Northumbrian OE (283).) Norsified English ‘arose at a time when Norse was still spoken but going out of use in its area’. The grammatical features in question include things already mentioned, such as the 3pl pronouns, *are* and *art* forms of the present tense of *be*, as well as a number of strong-verb forms, plural forms, comparative

(1971) in Thomason and Kaufman (1998: 360–4, notes 12–17). These remarks are largely corroborated by the discussion in Kastovsky (1992: 322–4), although he considers Sawyer’s estimates of the Danish population to be on the low side.

⁴ Thomason and Kaufman (1988: 270–2) provide their own characterization of the ‘ethnolinguistic areas’ of England and Scotland, which I am somewhat simplifying here.

forms, etc., along with the loss of some inflections (for example, the *ge-* prefix on perfect participles).⁵ The entire list is given in Thomason and Kaufman (1988: 293–5, Table 6). Many of these seem to be cases of contact-induced borrowing, whether due to direct contact or indirect contact (imperfect learning) in the sense defined above. The apparently rather rapid language shift suggests indirect contact, and this is what we described above. As we mentioned, this is compatible with the idea that word-order change was actuated by contact; but what militates against this is the complete lack of evidence that ON was in any way ‘more VO’ than OE.

One issue which is important is the inflectional simplification – notably in verbal inflection – which took place in ME, and which may have been very important for various syntactic changes (see the discussions in §2.1 and §3.4). Kroch and Taylor (1997: 317–18) suggest that the simplification of verbal inflection came about in Northern ME due to the effects of imperfect learning of English on the part of ON speakers. Thomason and Kaufman (1988: 277) argue that the simplification of the verbal inflection was a native, i.e. not a contact-induced, feature of Northumbrian OE, (although Traugott (1992: 177) says the loss of nominal inflection happened ‘possibly under Scandinavian influence’). They say that the ‘degree of simplification in the North, versus its rarity in the rest of the Danelaw, correlates with the rather high level of social upheaval prevalent in the North between 920 and 1100. It does not, however, correlate with anything in the structure of Norse’ (277). Furthermore, the dropping of final schwa in ME, which had the effect of eliminating the phonological exponence of many inflections and which is known to have taken place earlier in Northern than in Southern dialects of ME ‘cannot be blamed on the language contact situation’ (277) since it happened after 1250. It may be, then, that not all instances of morphological simplification that took place in ME, even if they took place first in the North, were caused by contact. (See Kastovsky (1992: 327–36) for a general discussion of the effects of contact between ON and OE.)

In conclusion, then, there is no question that contact with ON led to many changes in English. However, there is no strong basis for the idea that word-order change was among these. In this connection, it is worth

⁵ According to DeGraff (2005: 310), Schleicher (1850) had already suggested that the inflectional poverty of English as compared to Icelandic may have been due ‘to the much higher instance of language contact in the history of English’ (DeGraff 2005: 310).

recalling that change from OV to VO is quite common, and can take place independently of contact: the change in nineteenth-century Icelandic appears to be a case in point (see Hróarsdóttir (1999; 2000); Rögnvaldsson (1996)). The general account of markedness, and the formulation of the word-order parameter in (54F) of Chapter 3, indicate that VO is, all other things being equal, the unmarked setting for this parameter, and so this may be enough to explain the strong tendency for OV systems to change to VO that we noted in §1.6.2. The case for contact-driven change from OV to VO in OE and/or ME is not empirically clear, and is in any case not required: given the general similarity in word order between ON and OE, whatever internal factor might have caused ON to be ‘more VO’ than OE and thereby exert putative influence on OE could have been, and most likely was, operative in OE anyway. The evidence for competing grammars or formal optionality, depending on how one interprets the word-order variation in subordinate clauses, does not reveal anything about contact with ON, as it is present in varieties of West Saxon which could not have been influenced by ON.

5.2.3. *Substratum effects: Hiberno-English and Welsh English*

Let us now turn to substratum effects, which I will argue to be another case of indirect contact. The change to the system introduced through indirect contact is preserved once acquired by the younger group, following the schema in (1b). The usual notion of substratum in historical linguistics refers to the situation where a community gives up its original language in favour of a new one, but some feature or features of the original language survive and influence the structure of the adopted language. A standard example comes from the history of Spanish: Cantabrian Spanish is thought to have a Basque substrate, which may be responsible for the loss of initial /f/ in that variety, which spread to other dialects – notably Castilian – in the Middle Ages, giving, for example Spanish *hacer* from Latin *facere* (‘to do’) (Menéndez-Pidal 1982: 198ff.; Penny 1991: 79–82). This feature of Spanish may thus originate in language shift on the part of Basque speakers, who, as it were, spoke Spanish ‘with a Basque accent’ and therefore without initial /f/, since at this period it is thought that Basque had no /f/. (For a critical evaluation of this and other standard examples of substratum effects, see Hock and Joseph (1996: 382–5).)

Here I follow Thomason and Kaufman (1988: 116–8) in not distinguishing between the traditional notions of substratum, adstratum, and superstratum, although I will distinguish substrate from superstrate in the discussion of creoles in the next section; in that context, it is important to distinguish these notions in order to understand the ‘relexification’ account of creole genesis, as we shall see. All three notions, as traditionally used, refer to the socio-political relations between different groups shifting languages rather than to any structural features. French, for instance, is said to have a Celtic substratum since the inhabitants of Gaul who eventually abandoned Gaulish for Latin/Romance were conquered by the Romans, whilst it has a Germanic superstrate since the Franks conquered France somewhat later and then gave up their Germanic language for Gallo-Romance. Thomason and Kaufman (1988: 116) point out that one feature seems to correlate with the substratum vs. superstratum distinction: ‘superstratum interference is more likely to include lexical items than substratum interference is’. They also point out that French is in a superstratum relation with English, and this correlates with the large amount of lexical borrowing from French into English. One might propose something similar for ON in relation to English, given the discussion in the previous section.

The case of possible substratum influence I want to discuss concerns certain syntactic features of Hiberno-English, the English of Ireland, and Welsh English. In both Wales and Ireland, the inhabitants have been steadily shifting from Welsh and Irish respectively to English over several centuries, starting in the seventeenth century in the case of Irish (Ó Murchú 1993: 471–5) and the sixteenth in the case of Welsh (Owen Jones 1993: 536–63). The English spoken in these countries has certain features which may be attributable to substratum influence from the original languages, despite a paucity of loan words from them. Here I will look briefly at three features, one for each variety and one common to both.

Thomas (1994: 134ff.) presents several features of Welsh English which he attributes to substratal influence from Welsh. The first is ‘fronting’, shown in (3) (Thomas (1994: 137), his italics):

- (3) a. *Coal* they’re getting out, mostly.
b. *Singing* they were.
c. *Now* they’re going.

As Thomas says, these examples are equivalent to both clefting and pseudoclefting examples in Standard English, as the fronted material is new information:

- (4) a. What they're getting out mostly is coal.
 b. What they were doing was singing.
 c. It's now that they're going.

Welsh does not distinguish clefts from pseudoclefts, having instead a general fronting construction which fronts any XP to the left periphery of the clause, very likely to SpecCP, and can function to indicate new information. (See Tallerman (1996); Willis (1998); Roberts (2005) on this construction.) The following examples illustrate fronting of a DP, non-finite VP, and an adverbial PP:

- (5) a. [Y dynion] a werthodd y ci.
 the men Prt sold the dog
 'It's the men who have sold the dog.'
 (Tallerman 1996: 103)
- b. [Gadael y glwyd ar agor] a wnaeth y ffermwr.
 Keep the gate on open Prt did the farmer
 'Leave the gate open, the farmer did.'
 (Rouveret 1994: 77)
- c. [Ym Mangor] y siaradais i llynedd.
 in Bangor Prt spoke I last year
 'It was in Bangor I spoke last year.'
 (Tallerman 1996: 100)

Fronting of what appears to be a non-finite VP, as in (3b), is unacceptable in Standard English, whether to indicate new or old information. This variant of the construction at least, but perhaps the general use of fronting to indicate new information, may have entered this variety of English through imperfect learning of English by native speakers of Welsh and the subsequent transmission of this construction to later generations of native speakers of English, first perhaps bilingual in Welsh and then later not, following the schema for indirect contact in (1b). Although the precise grammatical analysis of this construction is unclear, it probably represents a parametric option connected to the structural realization of new information (and as such, its formal correlate concerns the distribution of EPP features on C).

The Hiberno-English example is the well-known perfect with *after*, as in (see Kallen (1994: 182 ff., 192); Cottell (2002)):

- (6) a. I'm after writing a letter.
b. She is after selling the boat.

This construction has a direct counterpart in Irish:

- (7) Tá sí tréis an bád a dhíol.
is she after the boat Asp sell
'She has sold the boat.'
(Kallen 1994: 192)

Once again, it is very tempting to see this construction as having originated in imperfect learning of English by native speakers of Irish, followed by its adoption into the native English of later generations and its retention after Irish had been abandoned, even as a second language. Again, the schema for indirect contact in (1b) applies. In this case, too, the precise nature of the parameter at work is not completely clear, although it is connected to the options for the realization of perfect *v*, and hence to the options we observed in Italian and various dialects in §4.1. (See Roberts (2005: 110–13) for some relevant discussion.)

Third, the construction common to both Irish and Welsh English is the option of inversion in indirect questions:⁶

⁶ Both of these examples illustrate what McCloskey (1992: 26 ff.) calls 'semi-questions' in the subordinate clause. The embedded clause does not convey true interrogative force, but rather a kind of indefinite proposition. McCloskey proposes a subtle semantic test which brings out this distinction; see the references given there as well as McCloskey's own discussion of the semantics of these clauses. The semantic difference between true questions and semi-questions is perhaps best intuitively seen in the distinction between *ask*, which takes a true question as its complement, and *ask about*, which takes a semi-question. Many varieties of English, including the variety of Hiberno-English described by McCloskey (1992) disallow inversion in subordinate semi-questions but allow it in true questions:

- (i) I asked them what would they do. (true question)
(ii) I asked them about what would they do. (semi-question)

Henry (1995: 107) explicitly states that the variety of Irish English she discusses differs from the one McCloskey investigated. It seems that the Welsh English reported by Thomas is like the variety Henry discusses rather than that discussed by McCloskey, although it is not easy to be sure on the basis of the evidence presented (cf. in particular, McCloskey's observation that in the variety he investigates, complements to *know* with inversion improve if *know* is negated).

- (8) a. I wouldn't know would there be any there now.
 (Welsh English, Thomas (1994: 138))
 b. I know is he going or not, but I'm not letting on.
 (Ulster English, Henry (1995: 105 ff.))

This is not allowed in Standard English, where inversion is only possible in main clauses and a small class of embedded declarative clauses. We saw in §1.3.1.2 that movement of the inflected auxiliary in T to C is blocked by the presence of a complementizer in C (see example (65)). In fact, the condition in Standard English is that any complement interrogative C, whether realized by a complementizer or *wh*-expression or not, blocks inversion, i.e. movement of T. In Hiberno-English and Welsh English, this condition does not seem to hold, as (8) shows. Both Welsh and Irish, as VSO languages, allow orders very close to those seen in (8). These examples are from Roberts (2005: 20–1):

- (9) a. Tybed a geith hi ddiwrnod rhydd wythnos nesa. (Welsh)
 I-wonder Prt will-get she day free week next
 'I wonder if she'll get a free day next week.'
 b. Chuir sé ceist ort an raibh tú sásta. (Irish)
 asked he question to-you Prt were you content
 'He asked whether you were content.'

However, there are good reasons to think that in fact the verb has not raised as far as C in (9), but only to T with the subject failing to raise to SpecTP. (See the discussion of Welsh in §1.3.1.1; that discussion carries over to Irish.) The element *a/an* glossed as 'Prt' in (9) is most likely to be an interrogative complementizer comparable to English *if*, a fact that would be consistent with the idea that the verb has moved only as far as T in these examples.

Nevertheless, the general proposal for transmission of these structures to later generations through imperfect learning of English on the part of speakers of Irish or Welsh can be maintained. Presumably native speakers of these VSO languages never learnt English so imperfectly as to attribute general VSO order to English, but they may have over-generalized inversion in embedded clauses, partly due to interference from the native VSO grammar due to imperfect adult L2 learning along the lines described in the previous section.⁷ This then becomes part of the PLD for subsequent

⁷ It is worth noting that Thomas (1994: 138) also comments that the 'elision of the conjunction [i.e. the complementizer, in the terminology being used here – IGR] (*if/whether*) is also facilitated by the Welsh rule of eliding the corresponding conjunction (*a/os*) in similar environments in the vernacular'. I do not know whether Irish allows a similar process of elision, however.

generations, and the substratum effect is once again created. The parameter in question concerns which kinds of C are able to trigger movement of T in embedded clauses; in terms of the discussion of T-to-C movement in §1.3.1.2, this is residual, symmetric V2.

The alternative to substrate analyses which is often put forward is that the same construction could develop independently of contact; in fact, we raised exactly this objection to the idea that ON actuated the OV-to-VO change in English above. In the case of (8), it would suffice to point to the numerous non-standard varieties of English which allow inversion in indirect questions, including the variety of Hiberno-English discussed in McCloskey (1992) and mentioned in note 6; AAVE (Mufwene 2001: 308); and other varieties of American English. However, as McCloskey (1992) and Henry (1995) have shown, there may be very subtle differences among varieties allowing inversion in embedded questions – see note 6. One might expect that the most liberal varieties are those which show substrate effects through indirect contact, in the sense described here. This must remain a question for further investigation, however.

5.2.4. A 'borrowing scale'

Let us finally consider Thomason and Kaufman's (1998: 74–6) borrowing scale, in relation to what we have said regarding contact and our general approach to syntactic change. (10) summarizes Thomason and Kaufman's scale, restricting attention to lexical and syntactic traits:

- (10) (1) Casual contact: lexical borrowing of content words only.
(2) Slightly more intense contact: lexical borrowing of some function words; syntactic borrowing of new functions and 'new orderings that cause little or no typological disruption' (74).
(3) More intense contact: function words, derivational affixes, inflectional affixes with borrowed vocabulary. Syntax: no complete typological change, but perhaps a partial one.
(4) Strong cultural pressure: moderate structural borrowing. '[F]airly extensive word order changes will occur ... borrowed inflectional affixes ... will be added to native words' (75).
(5) Very strong cultural pressure: heavy structural borrowing. 'Major structural features that cause significant typological disruption' (75).

Thomason and Kaufmann (1988: 97) state that the relation between ON and OE is of a different type, which they call typologically favoured borrowing; this is ‘structural borrowing at a higher level than the intensity of contact would seem to warrant, thanks to a close typological fit between source-language and borrowing-language structures’.

Aside from the first, all of these degrees of borrowing might involve the transmission of parameter values by disruption of the PLD according to the schemas for direct and indirect contact in (1). What distinguishes the degrees in (10, 2–5), to the extent that these distinctions are genuinely valid, is, on the one hand, the frequency of expression of a parameter: a change to a parameter associated with C, T, or *v* will have a greater effect than one associated with P, for example, since all clauses involve realizations of the former but not only some involve a realization of the latter. On the other hand, we could invoke something along the lines of Baker’s hierarchy of parameters in order to distinguish amongst (10, 3–5); the greater extent of borrowing might be connected to the relatively superordinate position of a parameter in the hierarchy. Clearly, a change in a more highly ranked parameter may have much more impact on the overall grammatical system, and thus require a much greater degree of contact for the younger group in the schema in (1) to be able to disregard the evidence for the very different indigenous structure. Something along the lines of Thomason and Kaufman’s scale could in fact be indicative of the parameter hierarchy and even be used as a tool to investigate it. This would represent a major, but very interesting, research project.

5.2.5. *Conclusion*

In this section I have considered how language contact can be looked at in relation to the general approach to syntactic change that I have been describing here. It can be integrated quite usefully into the approach and serve as a way of solving the Regress Problem, as well as giving us one way of seeing how Inertia can be violated. Here we looked at two cases of contact: contact between ON and OE in the Danelaw in the period 900–1200, and contact between Celtic languages and English in Ireland and Wales. We concluded that the OE-ON contact in the Danelaw is not necessary, and probably not sufficient, to explain the word-order changes in ME, contra Trips (2002). The possible Celtic substratum effects on English represent plausible accounts for some of the syntactic peculiarities

of the varieties of English in Ireland and Wales, although they may not be the only possible accounts. Both cases discussed here are probably indirect contact as in (1b). In §3.3 we saw an example of direct contact as in (1a): the ‘borrowing’ of preposition-stranding, along with some English prepositions, into Prince Edward Island French. This is a very straightforward case of lexical borrowing creating a new syntactic option, in this instance preposition-stranding.

In general, though, we can see that the evidence of language contact can be integrated into a parametric, acquisition-driven model of syntactic change quite unproblematically, and indeed in a way which yields up a number of interesting research questions.

5.3. Creoles and creolization

5.3.1. *Introduction: pidgins and creoles*

Having looked at contact in general in the previous section, let us now look at what has often been seen as a special or extreme case of contact: the development of pidgins and creoles. The study of pidgins and creoles has been an important part of linguistics since the pioneering work of Schuchardt (see the collected translations in Schuchardt (1979; 1980)). The essential interest of these varieties for historical linguistics lies in their origins. At least the external aspects of the origins of creoles are somewhat unusual, in that they arise from vernaculars developed in contact situations, usually known as pidgins. This has led to the idea that creoles might be significantly different from non-creole languages in that they may not have developed through the ‘normal’ mode of generation-to-generation transmission of language. Instead, they develop from pidgins, which are generally simplified communication systems limited to the contact situation in which they arise. Most importantly, pidgins are no-one’s native language. Holm (1988: 4–5) defines pidgins as follows:

A pidgin is a reduced language that results from extended contact between groups of people with no language in common; it evolves when they need some means of verbal communication, perhaps for trade, but no group learns the native language of any other group ... By definition the resulting pidgin is restricted to a very limited domain ... and it is no-one’s native language (emphasis in the original).

Creoles, on the other hand, are thought to arise when a pidgin functions as PLD for a generation of children, and thereby becomes a native language. To quote Holm (1988: 6) again:

A creole has a jargon or pidgin as its ancestry; it is spoken natively by an entire speech community, often one whose ancestors were displaced geographically so that their ties with their original language and sociocultural identity were partly broken. Such social conditions were often the result of slavery.

Thus the opposition between pidgins and creoles can be stated as follows: pidgins have no native speakers, are typically acquired by adults, have restricted communicative functions, are structurally simple, show inconsistent structural patterns, and may in fact not be compatible with the universal principles imposed by UG. Creoles, on the other hand, have native speakers, are acquired by children, have a full range of communicative functions, are structurally complex with consistent patterns just like any other language, and are compatible with what we know about universals, both from a typological point of view and in the sense of not showing features which violate UG.

If creoles arise exclusively from pidgins, as pidgins are described here, then this implies an ‘exceptional’ kind of origin for creoles. As already mentioned, this is the idea that has given rise to much of the interest in creoles. However, if the role of the superstrate or lexifier language (i.e. the – usually European – language that at least provides most of the pidgin and creole vocabulary) and that of the substrate language (i.e. the – most often African – language that was spoken by the displaced populations who are the ancestors of the eventual creole speakers) are taken into account, along with what is known about contact and substratum effects, then it is possible that the history of creoles does not feature such an exceptional ‘interruption’. This is the view advocated by Mufwene (1986; 2001) and DeGraff (2003; 2004; 2005; to appear), as we shall see.

5.3.2. The Language Bioprogram Hypothesis

It has frequently been pointed out that to the extent that the process of creolization involves, to quote Sankoff and Laberge (1973), ‘the acquisition of native speakers by a language’, creoles may be able to tell us much about language acquisition, language change, and learnability. Pidgins and creoles are therefore of central importance to the concerns of this book. Among the

best known work in this connection is that of Derek Bickerton (1981; 1984; 1999), who has argued that creoles give a direct insight into the language faculty. Bickerton (1981) put forward what he called the ‘Language Bioprogram Hypothesis’, whose central idea is that creoles are acquired on the basis of a radically impoverished trigger, so impoverished that the Language Bioprogram, the innate capacity which makes language and language acquisition possible, has a more direct relationship with the final-state system than in the case of non-creole languages. (In later work Bickerton (1984) identified this with Chomsky’s conception of UG, as I shall do here.) Non-creole languages are, as we argued in the Introduction to Chapter 1, underdetermined by experience, but the standard assumption is that aspects of experience profoundly influence the final state of language acquisition by cuing parameter values, etc. On the other hand, as we mentioned above, creoles have been taken by Bickerton and many others to have the characteristic property that their history features a break in the normal generation-to-generation ‘transmission of language.’ The special property of creoles is that they are based on highly defective, or even absent, PLD, as this comes from pidgin. In this situation, to quote Bickerton, ‘the human linguistic capacity is stretched to the uttermost’ (Bickerton 1981: 4). To further quote Bickerton:

It is debatable whether the P-input [pidgin input – IGR] is language at all ... P-input is in no sense a reduced or simplified version of some existing language: it is a pragmatic, asyntactic mode of communication using lexical (and very occasionally grammatical) items drawn mainly, but by no means exclusively, from the politically dominant language.

(Bickerton 1991: 365)

At the point of creole genesis, then, a new system is effectively ‘invented’. Because of this, Bickerton argues, creoles provide a unique window on the language faculty.

The really interesting claim in connection with the Language Bioprogram Hypothesis, which can be thought of as following from the idea that ‘the human linguistic capacity is stretched to the uttermost’ in creolization, is that, while creoles represent systems that conform to UG, they in fact only show a small amount of the variation that we know from the study of non-creole languages to be available. To put it in terms of principles and parameters, it seems that creoles occupy only a small sub-area of the general space of parametric variation made available by UG. (See §4.3.3 on the idea that the parameters of UG make available a multidimensional state space.) If this is so, then there is indeed something special about

creoles, and it is certainly natural to attribute this to the special circumstances surrounding their origins, given what we have just seen.

The evidence for the above claim comes from the striking morphosyntactic similarities that hold among creoles that are based on different lexifier languages and are widely dispersed both geographically and historically. Bickerton (1981, Chapter 2) lists twelve such properties, and other authors (D. Taylor 1971: 294; Muysken 1988: 289–92; Romaine 1988: 47–69) have listed others. It is most unlikely that these similarities are the result of historical borrowing or contact, and extremely unlikely that they are due to chance.⁸

Let us now look at some of the most striking morphosyntactic similarities that have been observed among creoles, and then consider how they might be understood in terms of some of the parameters of UG that we have been dealing with in this book. The first concerns the nature of verbal inflection. Holm (1988: 148) states that '[w]ith few exceptions, basilectal Atlantic creole verbs have no inflections'.⁹ Person–number agreement marking is entirely absent, including in creoles whose lexifier languages are inflectionally rich null-subject Romance languages such as Spanish and Portuguese. Tense, mood, and aspect are indicated by preverbal particles, whose nature we will return to below. Mufwene (1986: 134–6) shows that Kituba, a Bantu-based creole, lacks the typical Bantu pronominal prefixes on verbs indicating person and number and the agglutinating tense–aspect system. Person and number are expressed with overt pronouns, and tense–aspect by invariant particles. Verbal inflection is reduced to finite vs. non-finite marking and the passive, applicative, and causative extensions.

In our discussion of cues in Chapter 3, we suggested that the cue for V-to-T movement is person agreement in simple tenses (see Chapter 3,

⁸ The 'monogenesis' theory of creole origins holds that all creoles based on European languages are derived from Sabir, a Portuguese-based creole that was spoken in Africa in the fifteenth century, and which perhaps descended in turn from the older Mediterranean contact vernacular *Lingua Franca* (cf. Thompson (1961); Todd (1974); Whinnom (1965).) Mufwene (1986: 130–1) points out that the monogenesis theory only begs the question of universalist vs. substratist explanations for the nature of creoles, since we do not know how Sabir or *Lingua Franca* were formed; for a detailed and critical discussion of monogenesis, see Holm (1988: 44–52, 265–6). On *Lingua Franca*, see Holm (1988: 606–9). The table given in Romaine (1988: 89) indicates one possible set of 'monogenetic' relations.

⁹ Atlantic creoles are those spoken in 'the Caribbean area and coastal West Africa' (Holm (1988: 11)).

(54, B, c)). A corollary of the absence of the person–number agreement on verbs, then, is the absence of V-to-T movement in creoles. This property is particularly striking in French-based creoles, as the evidence for V-to-T movement is particularly clear in French. The following examples from Haitian Creole illustrate the lack of V-to-T movement in this language (see DeGraff (2005: 307–10), Roberts (1999: 304–7)):

(11) Adverb:

- a. Bouqui **repasse** déjà le linge. (French)
 Bouqui irons already the cloth
 ‘Bouqui is already ironing the clothes.’
- b. *Bouqui déjà **repasse** le linge.
- c. *Bouki **pase** déjà rad yo. (Haitian)
 Bouki iron already cloth the(ir)
- d. Bouki déjà **pase** rad yo.
 Bouki already iron cloth the(ir)
 ‘Bouki has already ironed the(ir) clothes.’

(12) Negation:

- a. *Jean ne pas **aime** Marie. (French)
- b. Jean n’**aime** pas Marie.
 ‘John does not love Mary.’
- c. Boukinèt pa **renmen** Bouki. (Haitian)
 Boukinèt NEG love Bouki
- d. *Boukinèt **renmen** pa Bouki.
 Boukinèt love NEG Bouki
 ‘Boukinèt does not love Bouki.’

Here we see that the usual diagnostics for V-to-T movement clearly show that Haitian has the ‘English’ value for this parameter. The same is true of the Indian Ocean creole Mauritian (from Green (1988: 459)):

- (13) li pa pu **dir** narjè
 he neg prt say nothing
 ‘He won’t say anything.’

Trinidad Creole French also shows this pattern (Hancock (1985), cited in Holm (1988: 378)).¹⁰

¹⁰ Réunionnais, the French-based ‘semi-creole’ (Holm 1988: 9–10, 392) spoken on the island of Réunion in the Indian Ocean, appears to have V-to-T movement:

- (i) Li **māz** pa sel.
 he eat not salt
 ‘He doesn’t eat salt.’

A second morphosyntactic similarity has to do with word order: creoles are almost without exception SVO (Bickerton 1981; 1984; 1988; Mühlhäusler 1986; Muysken 1988). There is nothing remarkable about this in creoles based on English or Romance, but it is striking that Dutch-based creoles such as Negerhollands and Berbice Dutch are also SVO in both main and embedded clauses (Holm 1988: 212). Similarly, Rabaul Creole German (also known as Unserdeutsch), spoken in Papua New Guinea, is consistently SVO; it is also not V2 (Romaine 1988: 30). SVO is of course a very common order among non-creoles, but SOV is just as common, and there is a significant minority of VSO languages. So creoles as a group can be distinguished from non-creoles in that they do not show non-SVO typologies; on the other hand, SVO itself is not confined to creoles. This is a good example of how creoles occupy just part of the space of variation that is attested in language in general.¹¹

However, Réunionnais is known to be more 'heavily influenced' by French than the other Indian Ocean creoles or Haitian. In fact, Réunionnais is explicitly excluded from the class of creoles by Bickerton (1981: 4). Baker and Corne (1982: 107) provide evidence that Réunionnais had greater superstrate contact than other French-based creoles. Another interesting intermediate case is Mesolectal Louisiana Creole, as discussed by Rottet (1993); DeGraff (1994); DeGraff and Dejean (1994); Roberts (1999). It has been claimed that Cape Verdean and Palenquero have V-to-T movement – see DeGraff (2005: 340–3) and the references given there.

¹¹ Some Indo-Portuguese creoles show at least optional OV orders. The variety spoken in Sri Lanka, reported in I. Smith (1978; 1979a, b; 1984) and discussed in Holm (1988: 288–90), shows some OV orders, along with 'case inflections on nouns, verbal inflections, postpositions, a phrase-final quotative particle and conditional marker, and various post-posed particles' (Holm 1988: 289); in other words, a considerable range of 'OV' properties. Holm further observes that all speakers of this variety are bilingual with Tamil, a language which, like the closely-related Malayalam (see Chapter 1, (129)), has these features. He suggests that 'it is possible that the creole's morphology is the result of recent wholesale borrowing associated with language death' (1988: 289).

There are other possible counterexamples to the claim made in the text. Bickerton (1988: 282) mentions Spanish-based creoles spoken in the Philippines which have VSO order; again, according to Bickerton, all speakers of these creoles speak a verb-initial Filipino language. Romaine (1988: 30–1) mentions Trader Navajo, which is VSO, along with Hiri Motu and Eskimo Trade Jargon, both SOV, but the latter is a contact jargon and the former a pidgin (Holm 1988: 584–7, 597–9; it is unclear whether Trader Navajo is a pidgin or a creole).

Consider next the null-subject parameter. It seems that creoles generally do not have referential null subjects. Here the interesting cases are creoles derived from null-subject languages such as Spanish and Portuguese. The following examples from the Spanish-based creole Papiamentu illustrate both the impossibility of omitting the subject pronoun (14b) and the impossibility of ‘free inversion’ (14c); see §1.2.1, on the relation between null subjects and free inversion. I have also indicated the corresponding grammatical Spanish sentences in parentheses (the examples are from Muysken (1988: 291)):

- (14) a. E ta kome.
 he ASP eat
 (él está comiendo)
 ‘He is eating.’
 b. *Ta kome.
 (está comiendo)
 ‘(S/he) is eating.’
 c. *Ta kome Maria.
 ASP eat Maria
 (está comiendo Maria)
 ‘Maria is eating.’

Nicholis (2004: 41–75) provides a very detailed survey of the status of the null-subject parameter in a wide range of creoles. He observes that referential null subjects are absent in the following creoles: Kriyol (Portuguese-based, spoken in Guinea-Bissau), Saramaccan (Spanish- and Portuguese-based, spoken in Suriname), and Cape Verdean (Portuguese-based, spoken in Cape Verde). Holm (1988: 202–3) states that ‘[i]t is not possible to omit the subject pronoun in Principe C[reole]P[ortuguese]’ (203) and that ‘the pronominal systems of the other Iberian-based creoles have some points in common with it’ (203), implying that unstressed subject pronouns are obligatory where null subjects can appear in the lexifier languages Spanish and Portuguese. It seems clear, then, that creoles in general do not permit referential null subjects.

The situation concerning non-referential null subjects is rather more complex. All of the creoles just mentioned allow these:

- (15) a. (A) (bi-) kendi/koto.
 it TNS hot/cold
 ‘(It) was hot/cold.’
 (Saramaccan; Byrne 1987: 76)
 b. Tawata jobe.
 PAST rain

- ‘(It) was raining.’
 (Papiamentu; Kouwenberg 1990: 46)
- c. Falta puku karu maja l.¹²
 lack little car hit him
 ‘A car nearly hit him.’
 (Kriyol; Kihm (1994: 48), cited in Nicholis (2004: 43))
- d. Sta faze kalor oji.
 is make hot today
 ‘It’s hot today.’
 (Cape Verdean; Nicholis 2004–5: 73)

Moreover, many creoles whose lexifiers are non-null-subject languages show the same pattern in allowing non-referential null subjects but disallowing referential ones. This is true of Berbice Dutch (Dutch-based), Haitian (French-based), Jamaican (English-based), and Mauritian (French-based), as the following examples show:

- (16) a. Te fè frèt.
 ANT make cold
 ‘It was cold.’
 (Haitian; DeGraff 1993: 72)
- b. O bi masi mɛnɛ dunggrə.
 3sg say must middle night
 ‘He said (it) must be midnight.’
 (Berbice Dutch; Kouwenberg (1994), cited in Nicholis (2004: 48))
- c. (I) look like im nuh like yu.
 (EXPL) look like 3sg NEG like 2sg
 ‘It looks like s/he does not like you.’
 (Jamaican; Durreleman (2004), cited in Nicholis (2004: 64))
- d. Posib Pyer lakaz.¹³
 possible Peter house

¹² Kriyol seems to differ from the other creoles mentioned here in that it does not allow ‘meteorological’ null subjects:

- (i) I na burfa.
 ‘It’s drizzling.’
 (Nicholis 2004: 44)

¹³ Mauritian also allows an indefinite, generic null subject:

- (i) Fer rom ar disik.
 make rum from sugar
 ‘Rum is made from sugar.’
 (Syea (1992), cited in Nicholis (2004: 67))

‘It is possible that Peter is at home.’
(Mauritian; Nicholis 2004–5: 69)

The creole evidence thus seems to confirm Rizzi’s (1982: 143) proposal that the null-subject parameter should be seen as two ‘related but autonomous parameters’, one of which determines the availability of null pronouns, the other determining the possibility of null referential pronouns. On this view, we can conclude that the former parameter has the positive value and the latter the negative one in creoles: hence non-referential null subjects are allowed and referential ones are not. Creoles seem to pattern this way across the board, irrespective of the status of their lexifier languages with respect to the null-subject parameter.

A fourth property which creoles seem to share is the absence of ‘special’ complement clitics, typically occupying a preverbal position in finite clauses. Such clitics are found in all the Romance languages, but appear to be systematically absent in Romance-based creoles. Holm (1988: 211) comments that ‘object pronouns ... always follow the verb in Romance-based creoles’. The following contrasts between French and Haitian illustrate the situation:

- (17) a. Bouqui l’aime. (French)
 Bouqui 3sg-like
 b. Bouki renmen li. (Haitian)
 Bouki like 3sg
 ‘Bouqui likes him/her/it.’
 c. *Bouki li renmen.

Haitian is quite typical of Romance-based creoles in this respect.

A fifth morphosyntactic property is the nature of the preverbal tense/mood/aspect (TMA) particles. These are invariant, monomorphemic elements which appear in a fixed sequence preceding and adjacent to the first verb. Some particles are postverbal in some creoles (for example, the Cape Verdean anterior marker *ba* (Holm 1988: 149)), but the overwhelming tendency is for preverbal positioning of these elements. They constitute the basic way in which tense, mood, and aspect are expressed in creoles, and are common to all creoles. These particles have appeared in several of the examples given above, for example, the Haitian anterior marker *te* in (16a), the Papiamentu past marker *tawata* in (15b), and the progressive marker *ta* in (14a). Here are some further examples from Green (1988: 453):

- (18) a. Li **pa** **ti** kapav fer sa. (Mauritian)
 he NEG ANT able do that
 ‘He couldn’t do that.’
- b. Mo **te** bezwẽ fε l. (Haitian)
 I ANT must do it
 ‘I had to do it.’
- c. E **tabata** sigi bende piska. (Papiamentu)
 he PAST CONT sell fish
 ‘He went on selling fish.’

The TMA particles cannot be fronted with the fronted verb in the ‘predicate-cleft’ construction, which involves focusing and fronting of a copy of the VP. This is illustrated by the following contrast in Papiamentu:

- (19) a. Ta [ganja] Wanchu a ganjabo.
 FOC lie John ASP lie-you
 ‘John has really lied to you.’
- b. *Ta [a ganja] Wanchu a ganjabo.

Negation (*pa* in both Haitian and Mauritian) must precede all TMA markers:¹⁴

- (20) a. Jan **pa** t ava ale nan mache. (Haitian)
 Jan NEG ANT MOOD go in market
 ‘John would not have gone to the market.’
- b. Jan **te** (***pa**) ava (***pa**) ale (***pa**) nan mache.
 Jan ANT NEG MOOD NEG go NEG in market

See also the Mauritian example in (18a), where *pa* precedes the anterior marker *ti*.

The TMA markers arise through fairly typical ‘grammaticalization paths’, as the following quotation from DeGraff (1993: 75) illustrates (see also Mufwene (2001: 54–6, 77)):

Pral, marking future, also means ‘to go’; *dwe*, marking obligation or possibility, also means ‘to owe’; *fini*, marking completion, also means ‘to finish’; *konnen*, marking habituality, also means ‘to know’; *sòti*, marking recent past, also means ‘to leave’; etc.

There is also a more detailed discussion of the origins of the Haitian TMA markers in DeGraff (2005: 322–3).

¹⁴ This is not true in all French-based creoles. DeGraff (2005: 356) points out that in Louisiana Creole *pa* follows the anterior marker *te*.

All the above points clearly indicate that the TMA markers should be analysed as functional elements occupying T or v in terms of the basic clause structure I have been assuming here. The predicate-cleft evidence suggests that they are VP-external, and their preverbal, post-subject position clearly points to T or v. The fact that they may correspond to grammaticalized main verbs in the lexifier languages, combined with the loss of V-to-T movement where the lexifier had this, suggests that these elements have a diachronic source which is comparable to that of the English modals as discussed in Lightfoot (1979; 1999); Roberts (1985; 1993a); Warner (1993); Roberts and Roussou (2003); and §2.1. See also Mufwene (2001: 55).

The morphosyntactic properties discussed above and exemplified in (11)–(20) illustrate some of the striking similarities among creoles. Others have also been suggested, such as serial verbs (Holm 1988: 183 ff.); negative concord (Bickerton 1981: 65; Holm 1988: 171–3; Déprez 1999; 2000); lack of inversion in direct questions (Bickerton 1981: 70; Holm 1988: 212; DeGraff 1993: 75); various complexities in copular constructions (Holm 1988: 174 ff.; DeGraff 1992; 1995); overt wh-movement (Bickerton 1988: 282); and the absence of passives (Bickerton 1981: 71; Green 1988: 453), although here it is argued that Crioulo is an exception. What we consistently observe is that the range of parametric variation attested in creoles is a small subset of the variation we know to be available in UG. There are no attested examples of SOV creoles (but cf. note 11), ergative creoles, V2 creoles, creoles allowing referential null subjects, creoles with rich fusional morphology, etc. If this is genuinely the case, then some kind of explanation is required.

It is important at this stage to be clear on two points. First, it is not being claimed that creoles are all identical in their syntax: Muysken (1988: 291–3) observes a number of differences among serial-verb constructions in creoles, as well as pointing out (1988: 294) that some creoles allow Preposition-stranding while others do not; we have observed in notes 12 and 13 above that there are certain differences among creoles regarding null subjects, although the generalization that no creole allows fully referential null subjects appears to hold; finally, there appear to be a number of detailed and intricate differences among the systems of TMA markers found in creoles, despite the general similarity in the existence of such systems in the first place (see Holm (1988: 148 ff.); Muysken (1988: 291)). Second, it is not being claimed that creoles are synchronically exceptional in their morphosyntax.

None of the properties discussed above is absent in non-creoles: NE and the Mainland Scandinavian languages lack V-to-T movement (see §1.3.1.1); German and Icelandic have both been analysed as allowing expletive but not argumental null subjects (see (13) of §1.2.1); the North Germanic languages and English lack ‘special’ pronominal clitics, and, as we already mentioned, the English auxiliary system is rather like a system of TMA markers. What *is* being suggested is that there is a striking overall typological similarity among creoles, in that they appear to occupy a rather small space of the overall variation made possible by UG.

As mentioned, Bickerton (1981) explained this kind of observation by invoking the Language Bioprogram. This forces grammars to take on a certain form under the extreme conditions of acquisition based on pidgin PLD. In this way, it may emerge that creoles can provide a special kind of ‘window’ onto UG.

This account has some appeal, but it raises a problem as it stands. The principles-and-parameters approach cannot treat some set C of languages as ‘closer’ to UG than its complement set C’. We must maintain that creoles have exactly the same relationship to UG as any non-creole language, in that they represent a system of principles and parameters with the values of the parameters fixed (see also Lightfoot (1991: 182), DeGraff (2005: 343)). However, what we could maintain, more along the lines explored in Bickerton (1984; 1988), on the basis of the idea that the PLD is particularly deficient in the case of creole acquisition, is that creoles can tell us something about the default, unmarked values of parameters. This idea is put forward in the following terms by Bickerton (1988: 282):

The consistency of this typology, despite the absence of any consistent empirical model for it, argues strongly that in addition to universal principles of syntax we must assume the existence of an unmarked set of grammatical options by which those principles can be realized.

In the discussion of markedness and parameter setting in §3.4 and §3.5, I suggested that in the absence of a clear expression of the value of a given parameter, the default option is always taken. The absence of a clear expression of a value for a parameter amounts to weak P-ambiguity, in the sense defined in (21c) of Chapter 3, and repeated here:

- (21) A weakly P-ambiguous string expresses neither value of p_i and therefore triggers neither value of p_i .

It seems natural to think that much pidgin input may be weakly P-ambiguous in this sense, given what is known about the nature of pidgins (and cf. the quotations at the beginning of this section from Holm and Bickerton). In that case, as Bickerton suggests, it may be that owing to the nature of the pidgin input, creoles demonstrate a high preponderance of unmarked parameter-settings and tend to look alike as they all tend to have the same default parameter values. This view does not imply that creoles are qualitatively different from non-creoles, since unmarked parameter-settings are equally available to non-creoles.

In (54) of Chapter 3, I proposed what the default values for the six parameters discussed in Chapters 1 and 2 might be. Three of these, the null-subject parameter, the V-to-T parameter, and the head parameter(s), are relevant here, given the above discussion of the properties of creoles. They are repeated here:

- (22) A. *Null subjects*
- a. *Parameter*: Finite T {has/does not have} sufficient specification of agreement features φ to bear the subject thematic role/Agree with *pro* in SpecTP.
 - b. *Default*: φ is absent.
 - c. *Cue/expression*: ‘rich’ agreement morphology on T- and/or V-elements.
- B. *V-to-T movement*
- a. *Parameter*: Finite T {has/does not have} an EPP feature which attracts V.
 - b. *Default*: EPP is absent.
 - c. *Cue/expression*: (finite) V is marked with person agreement in all simple tenses.
- F. *The head parameter(s)*
- a. *Parameter*: a head H {has/does not have} an EPP feature triggering movement of its complement to its specifier.
 - b. *Default*: EPP is absent.
 - c. *Cue/expression*: overt complement>head orders.

It is clear that the creole value of the parameter is the default one in each case: non-null subject (for referential null subjects), no V-to-T, and head-initial. This corresponds to the lack of expression of the relevant inflectional morphology in the first two cases and to the lack of relevant word orders in the PLD in the third. This reasoning can be extended to the other properties characteristic of creoles discussed above: the absence of complement clitics and the presence of TMA markers. Although I have not stated what the ‘clitic parameter’ might be, and do not intend to delve in detail

into this tricky question here, it is very likely that complement clitics are moved to their preverbal position, as was originally argued by Kayne (1975) and in a great deal of subsequent work. If this is so, then, following the general characterization of markedness put forward in §3.4, which entails that parameter values associated with movement are more marked than those which are not, the presence of complement clitics represents a marked option relative to their absence. Hence once again creoles show an unmarked parameter setting in systematically lacking special pronominal clitics. Finally, TMA markers, as grammaticalized main verbs or auxiliaries merged as T or v, represent an unmarked option in relation to the expression of tense, mood, and aspect through inflections combined with either Agree or V-movement.

So we see that Bickerton's conjecture as summarized in the above quotation has some support. Once again, it is worth emphasizing that this does not entail that creoles are exceptional. As DeGraff (2005) points out, non-creoles may also develop the unmarked parameter values we have observed to hold in creoles. In fact, as I pointed out in Roberts (1999: 317), English has all the properties that we have attributed to creoles, and, assuming that Proto-Germanic was a null-subject language, all the relevant parameters have changed in the history of the language. This does not imply that English has undergone creolization; this idea is criticized at length in Thomason and Kaufman (1988: 263–332), as we mentioned in the previous section. Nor does it imply that creoles are special; it simply shows that 'normal' processes of change can lead to a series of unmarked parameter values. This may be more common in creoles owing to the fact that PLD consisting of pidgin is relatively prone to weak P-ambiguity. We will come back to this point when we discuss DeGraff's recent work (DeGraff 2003; 2004; 2005; to appear) below.

5.3.3. *The substratum/relexification hypothesis*

A different and widespread view of the nature of creoles, which also purports to explain the kinds of morphosyntactic similarities among them that we have observed, is the substratum hypothesis. The central idea here is that the shared grammatical features of creoles derive from the language or languages originally spoken by the creole speech community, along with relexification from the superstratum language. At least informally, one can

think of this as substituting words from the lexifier languages into syntactic structures characteristic of the substratum language(s). Given the discussion of substrate effects in terms of imperfect learning in the previous section, the idea would be that creoles arise through an extreme case of imperfect learning of the superstrate, lexifier language, such that the syntax of the substratum language is preserved to a very large extent. No doubt such extreme imperfect learning can be attributed to the extralinguistic situation in which pidgins are formed, which was certainly not one in which pidgin speakers attempting to learn the lexifier language as a second language would have been given any encouragement or tuition. The system with the syntax of the substratum language and the lexicon of the superstrate language then forms the PLD for subsequent generations, and in this way the substratum may be maintained for many generations.

A number of researchers have suggested that some of the features that we mentioned above, which Bickerton and others have taken as evidence for something like a language bioprogram, or the prevalence of unmarked parameter values in creoles owing to the circumstances of creolization, are in fact evidence of a substratum. In the Atlantic creoles, at least, this substratum is usually taken to be West African; the people who were forced into slavery in the Caribbean, Latin America and elsewhere originated in West Africa. This view has been put forward by various researchers, and arguably goes back to Schuchardt (see the discussion in Holm (1988: 27–35)). More recently, it has been put forward by D. Taylor (1977); Boretzky (1983); Muysken (1988); Koopman (1984; 1986); Lefebvre and Lumsden (1989); Lefebvre (1998); and Lumsden (1999). Taylor summarizes the basic ‘substratist’ position, saying ‘[w]hile African loan words are relatively few in most West Indian creoles . . . African loan constructions are both common and striking’ (D. Taylor (1977: 7), quoted in Holm (1988: 65)).

One striking property is the predicate-cleft, or verb-topicalization, construction illustrated in (19) above. This construction does not exist in English or French (although something very like it exists in Colloquial Italian) and is not an obvious candidate for an unmarked parameter setting, at the very least since it appears to involve the copying component of the Move operation; see Box 1.1 of Chapter 1 on some technical aspects of Move. It is, however, quite widespread in creoles (Romaine (1988: 104) gives examples from Sranan, Krio, and Mauritian), and is also found in some languages of West Africa, as the following examples (again from Romaine (1988: 104), with my parentheses; see the sources cited there) show:

- (23) a. [Mi mun] ni won mun mi. (Yoruba)
 me take is they took me
 ‘They actually arrested me.’
- b. [Hwe] na kwasi hwe ase. (Twi)
 fall is Kwasi fell down
 ‘Kwasi actually fell.’

The Kwa languages, spoken in several West African countries (Ivory Coast, Ghana, Togo, and Benin), are often posited as potential substratum languages, at least for Atlantic creoles. Among the properties of creoles that might be attributable to a West African substrate, such as the Kwa languages, are SVO order, preverbal TMA particles, serial verbs, the form of comparative constructions, and postnominal determiners.

Regarding SVO order, according to the data in Haspelmath *et al.* (2005), ten out of eighteen Kwa languages surveyed are SVO, with no relevant information available on the others. On the other hand, Bambara, another West African language which is sometimes mentioned as a possible substrate language (see Holm (1988: 149)), is OV. Some of these languages show a complication to basic VO order, however, in that in compound tenses a non-pronominal object precedes the main verb but follows the auxiliary. The following examples, from DeGraff (2005: 305, and see the sources cited there), illustrate this:

- (24) a. Ûn Dú mŌlinkún. (Fongbe)
 I eat rice
 ‘I eat rice.’
- b. Ûn Dò mŌlinkún Dú wε. (Fongbe)
 I be rice eat Prt
 ‘I am eating rice.’

Moreover, Aboh (1999; 2005) analyses the Gbe languages – including Fongbe – as having systematic V-to-T movement. Haitian and the other Atlantic creoles entirely lack this kind of alternation. As DeGraff (2005: 306) points out, this ‘is unexpected in the strict-relexification proposals’.

Regarding TMA particles, there are certainly some striking similarities between creoles and Kwa and other West African languages, clearly exemplified in the table in Holm (1988: 149). However, the data in Haspelmath *et al.* (2005) is rather equivocal on this point: only three of the Kwa languages have data on tense–aspect marking, and of those two have preverbal particles and one, Ewe, has an inflectional future and no other tense marking at all. Yoruba has no tense marking and Bambara has

suffixes. DeGraff (2005: 304) points out that postverbal and suffixal aspectual markers have been documented in the Gbe languages (see also Aboh (1999; 2005)). Once again, then, the case for substratum effects is not straightforward.

We mentioned Baker's Serial Verb Parameter in §3.5.4. In languages which allow serial verbs, more than one verb can appear in a single VP (or perhaps vP; Baker 2001: 141). We gave an example there from the West African language Edo, which I repeat here:

- (25) Òzó ghá lè èvbàré khiẹ'n.
 Ozo will cook food sell
 'Ozo will cook the food and sell it.'
 (Baker 2001: 140)

Verb serialization is a widespread property in the world's languages, being found in Chinese and other East Asian languages, and many African languages, including the Kwa languages and other West African languages. As such, it may be a good candidate for substrate influence.¹⁵

A common form for comparative constructions in creoles is a serial-verb construction with a verb meaning 'exceed' following the expression of the standard of comparison and the adjective. Holm (1988: 188) gives examples of this from Principe Creole Portuguese, Lesser Antillean (French-based), Ndjuka, and Gullah (both English-based). Here is the Ndjuka example:

- (26) A bigi pasa mi.
 'He is taller than I.'
 (Hancock (1979: 12), cited in Holm (1988: 188))

According to the data in Haspelmath *et al.* (2005), this kind of comparative is very common in West Africa, being found in Yoruba, Igbo, and Hausa, for example, although no data is available concerning the Kwa languages. This, then, is another good candidate for a substratum effect, although it seems very likely that this kind of comparative construction is connected to the existence of verb serialization.

¹⁵ One point worth noting stems from our discussion of Baker's parameter hierarchy in §3.5.4: there we pointed out that Baker's hierarchy makes the prediction that, if a language loses V-to-T movement, then it simultaneously loses the possibility of having VSO order or null subjects, but may go on to develop serial verbs. We mentioned there that English- and Romance-based creoles support this prediction, in that they lack V-to-T movement and (referential) null subjects, and have both SVO order and serial verbs. We have now seen the evidence for this.

Finally, a striking common property of many creoles and at least some West African languages is the presence of postnominal articles (Holm 1988: 190–2; Lumsden 1999). The following examples illustrate the similarities between the creoles Haitian and Principe and the West African languages Fongbe and Yoruba:

- (27) a. tab la
 table the
 ‘the table’
 (Haitian; Lumsden 1999: 145)
- b. bato a yo
 boat the pl
 ‘the boats’
 (Haitian; Lumsden 1999: 145)
- (28) a. DiDè Ò IÈ
 sketch the pl
 ‘the sketches’
 (Fongbe; Lumsden 1999: 147)
- b. wèamá Ò IÈ
 book the pl
 ‘the books’
 (Fongbe; Lumsden 1999: 147)
- (29) a. básta di óru sé
 can of gold the
 ‘the can of gold’
 (Principe; Boretzky (1983: 97), cited in Holm (1988: 190))
- b. owó tí nwọ fún mi náà
 money which they gave me the
 ‘the money which they gave me’
 (Yoruba; Rowlands (1969: 197), cited in Holm (1988: 190))

These examples show that the article comes after the noun and various adnominal complements and modifiers, such as relative clauses, but, at least in Haitian and Fongbe, precedes the number marker. According to the data in Haspelmath *et al.* (2005), this pattern is found in the Kwa languages Akan, Ewe, and Gã (in addition to Fongbe), as well as in Bambara. This looks like a further good candidate for a substrate feature, then.

The substratum idea is most strongly supported by constructions which do not obviously represent unmarked parameter settings, which are not cross-linguistically common, which are robustly attested both in West African languages and in creoles, and which are not found in the superstrate, lexifier languages. Of the examples just reviewed, postnominal articles and

predicate-clefting appear to meet these criteria. SVO order and TMA particles, aside from not being very robustly attested in West African languages and, in the former case, being common in the lexifier languages, may represent unmarked parameter values. Further, serial verbs may represent an unmarked parameter value, in that material is merged as *v* rather than *V* moving there (*pace* Nylander 1986), although they are also a feature of the substrate but not the superstrate. It seems, then, that there is something of a case for a substratum explanation for at least some properties of creoles.

We could therefore suppose that speakers of Kwa languages carried over the syntactic properties of those languages in their attempts to learn French, English, Dutch, Spanish, or Portuguese, for the most part simply substituting lexical items from the European languages into their native syntactic structures. This relexified Kwa would have constituted the PLD for subsequent generations, and the substratum would have thereby persisted across the generations, much along the lines of the general scenario for substratum effects that I suggested in the previous section.

However, simple relexification cannot be the whole story, as DeGraff points out in his discussion of verb- and object-placement in Haitian and Fongbe (see the quotation after (24) above). One feature which appears to be widespread in the Kwa languages, according to Haspelmath *et al.* (2005), is postpositions. But, with the notable exception of Berbice Dutch (Holm 1988: 210), creoles are typically prepositional. Arguably, then, the substratum hypothesis needs to be supplemented with an account of which features of the substrate languages are most likely to be retained, and why. And here again markedness theory may have a role to play.

Aside from the empirical issues discussed above, we can identify two conceptual problems with the relexification approach. First, if parameters are associated with lexical entries, as we are assuming here, then it is difficult to see how the notion of relexification can be formulated. The intuitive idea that lexical items from one system are inserted into the syntactic structures of another cannot be maintained under the minimalist assumptions being adopted here. In fact, the technical approach to lexical insertion assumed in Chomsky (1995, Chapter Four) and subsequent work does not allow for lexical items to be substituted into slots created by syntax; instead, merging lexical items creates syntactic structure – this was implicit in our discussion of Merge in the Introduction. The second problem is very clearly articulated by DeGraff (2005: 299) in the following passage, where he is discussing Lefebvre's (1998) claim that the creation of

Haitian involved relexification of West African substrate with French vocabulary:

Lefebvre must assume that the Creole creator was somehow able to segment and (re)analyze French strings and adopt and adapt a great deal of French phonetics and surface order – down to the phonetic shapes and surface distribution of many affixes and grammatical morphemes – while ignoring virtually *all* abstract structural properties of French. Such a feat would make the Creole creator unlike any other language learner documented in the psycholinguistics and language-acquisition literature. After all, word segmentation and word- and affix-order are reflexes of *abstract* morphosyntactic properties.

(DeGraff 2005: 299, emphasis in original)

The evidence for Very Early Parameter Setting which we reviewed in §3.1 bears out DeGraff's point: the evidence is that, if anything, parameters are set before many surface forms are in fact acquired. If this is so, then the question why the French grammatical system was not acquired along with its lexicon by the creators of Haitian becomes very acute. DeGraff (to appear: 28) also points out that this implies that creoles cannot be entirely new creations, as implied by the Language Bioprogram Hypothesis, since they are made up of lexical items from the superstrate language which presumably retain some of the associated abstract morphosyntactic properties. The idea that the lexical items of an alien system come associated with certain parameters is also supported by our discussion of PEI French in §3.3.1. There we saw that Preposition-stranding was introduced into this variety along with prepositions borrowed from English, and then generalized to native French prepositions.

5.3.4. Conclusion: how 'exceptional' are creoles?

In conclusion, the case that creoles typically favour unmarked values of parameters, as discussed by Bickerton (1984; 1988: 282; 1999: 56ff.), has a certain amount of evidence in its favour, as we have seen. At the same time, there is some evidence for substrate effects from West African languages, although this is not always straightforward. Some features which seem to be widespread in creoles could perhaps be equally well accounted for either way; this may be true for SVO order, TMA markers, or serial verbs, for example. But in fact, as Mufwene (1986: 129–30) pointed out, there is no real contradiction between the 'universalist', markedness-based view of

creoles and a substratum account of the origin of at least some common traits: ‘most of the features of pidgins and creoles that the substrate hypothesis has been claimed to explain are not really accounted for unless some universal principles are accepted to apply at some stage in the formation of these languages’. It is clear that there is nothing about the theory of markedness that would prevent the Kwa languages, or any other group, having a number of parameters set to unmarked values: indeed at least VO order and perhaps serial verbs (with the corollary of lack of V-to-T movement if Baker’s parameter hierarchy as discussed in §3.5.4, and above, is right) represent unmarked parameter settings on our assumptions. These parameter settings would naturally emerge in the creoles, either through the PLD or by default. More marked parameter values, such as postpositions perhaps, may not be sufficiently triggered by the pidgin PLD, although some, for example, postnominal articles, must be, given what we saw above. As MacMahon (1994: 280) puts it, ‘[w]hen structures in different substrates coincide, these will be especially likely to be introduced into the creole; the bioprogram is here seen as a last-resort explanation, to be invoked when the relevant substrate structures conflict or no evidence for a particular structure is available’. In other words, markedness considerations become crucial when the PLD is either strongly or weakly P-ambiguous.

This brings us to a final point on the topic of creoles. As I stated at the beginning of this section, much of the interest in pidgins and creoles has been stimulated by the idea that these systems can tell us something special about the nature of language change, language acquisition, or UG. This is because the process of creolization has been thought to represent a break in the usual generation-to-generation transmission of language. In recent papers, however, DeGraff (2003; 2004; 2005; to appear) has argued against what he calls ‘creole exceptionalism’ of this kind. His view is that creoles have emerged through normal processes of language change, a view he supports by pointing out that observed changes in the history of French, the Scandinavian languages, and, in particular, English, have yielded in many cases similar results. (We observed above that English has changed most of the parameters discussed above in connection with creoles, and in the same direction, in the course of its history.) As he says, ‘H[aitian]-C[reole] morphosyntax does not, *and could not*, isolate HC and its diachrony in some exclusively “Creole” empirical domain’ (2005: 314). This view may well be correct: the only thing which may be in any way ‘unusual’

about creoles is that at the point of creolization weak P-ambiguity may have been more prevalent in the PLD, given the nature of pidgin, than is usual. ‘Grammatical inventions’, to use a term introduced in Rizzi (1999), arise through the usual process of creation of a grammar by language acquirers on the basis of whatever PLD is available, but are perhaps more likely to arise when the PLD is prevalently weakly P-ambiguous. DeGraff (2005: 317) considers creole genesis to involve imperfect learning by adults in a ‘learner-unfriendly’ environment, with the subsequent consequences for the PLD of later generations, rather than specifically to a pidgin. He suggests that, once this is taken into consideration, creole genesis is really a further case of imperfect second-language learning. If DeGraff is right, then creole genesis may not be qualitatively different from the cases of imperfect learning discussed in the previous section. However, even this approach begs the question of how much weakly P-ambiguous PLD is usual, and this is a question that simply cannot be answered in our current state of ignorance regarding the relation between (first or second) language acquisition and language change. Lightfoot (2006: 139ff.) reaches a view on creolization largely compatible with DeGraff’s views, while a dissenting view is expressed in Bickerton (2004).

To conclude, I would like to quote my own comment in Roberts (1999: 317) on the allegedly exceptional nature of creoles and what this might be able to tell us about UG and language acquisition:

What gives us a privileged view of UG, and of the nature of the parameter-setting algorithm, is not creoles but language change. Creoles are particularly interesting in that they represent an extreme of language change, but it is the mechanisms of language change, which are ubiquitous in the history of every language and every language family, that have made creoles what they are.

To this I would only add that the ‘extreme’ aspect of language change in the case of creoles may in fact have more to do with the extreme social conditions under which these varieties were formed than with any intrinsic aspect of the linguistic mechanisms of change. The parameter-setting device may have functioned quite normally, but under abnormal external conditions, in this case.

5.4. Language creation in Nicaragua

In this section I want to turn to what might really be an exceptional case: language creation in Nicaragua. I will summarize some very interesting and

important research on **sign language** in Nicaragua, reported in Kegl, Senghas, and Coppola (1999). (See also Senghas (1995a, b), Kegl (to appear).) This appears to be a case of language creation in that a new signed language has emerged where none existed before. If they are correct, Kegl, Senghas, and Coppola's (henceforth KSC) work has implications for language-acquisition theory, the study of pidgins and creoles, and for the study of language change under the general assumptions I have been arguing for here: the creation of a language is in a sense an extreme case of the reanalysis of PLD that underlies change, and so this case may have something to tell us about language change in general.

Before embarking on the discussion of signed languages in Nicaragua, a word or two regarding sign languages in general is perhaps in order. It is now an accepted result of modern linguistics that sign languages as used by Deaf communities¹⁶ in many parts of the world are true languages in every sense (see Goldin-Meadow (2005: 201–2)). They differ from spoken languages only in the modality of transmission: gestural/visual as opposed to oral/aural. The signed languages which have been studied show all the structural features of spoken languages, including notably a syntax which has all the hallmarks of a generative system, being discrete, algorithmic, recursive, and purely formal. Signed languages also have a phonology, in that signs, which have a meaning, are made up of smaller units which themselves lack meaning, just as words (or morphemes) in spoken languages are made up of phonemes. Phonological processes such as assimilation have been observed, and phonological units such as the syllable proposed (see Sandler and Lillo-Martin (2001: 539–42) for a summary of the evidence for sign-language phonology). And signed languages show the full range of semantic properties of spoken language. To quote Sandler and Lillo-Martin (2001: 534), sign languages:

are natural languages, in the sense that they are not consciously invented by anyone, but, rather, develop spontaneously wherever deaf people have the opportunity to congregate and communicate regularly with each other. Sign languages are not derived from spoken languages; they have their own independent vocabularies and their own grammatical structures. Although there do exist contrived sign systems that are based on spoken languages . . . , such systems are not natural languages.

¹⁶ I follow the standard practice in work on sign languages in using the capitalized term 'Deaf' to refer to Deaf communities and their members, and the non-capitalized 'deaf' to refer to hearing loss.

The emergence of a new sign language, then, is an instance of the emergence of a new natural language, and as such of great interest. Let us now look at KSC's documentation of this remarkable event. These developments are also summarized and discussed in Lightfoot (2006: 152ff.).

KSC look at the signing of a community of approximately 500 Deaf children and young adults in Managua, Nicaragua. Their claim is that the sign language this community uses for internal communication has come into existence since approximately 1980. Writing in the mid-1990s, they state that this 'newly emergent language has been in existence for barely more than a decade' (178). They therefore have 'one of the first documented cases of the birth of a natural human language' (178).

The background to this situation lies in the social and political situation of Nicaragua. Until the Sandinista revolution in 1979, Nicaragua was ruled by a dictatorship which made care for the disadvantaged and education for the majority of the population a low priority. Accordingly, there were no schools or any kind of social or welfare provision for the Deaf. Furthermore, there was considerable social stigma attached to deafness. For these reasons, there was no Deaf community in pre-Sandinista Nicaragua, and correspondingly no sign language. In 1980, the Sandinista government set up a number of special schools as part of its general campaign to provide education for the population as a whole. In Managua, about 500 Deaf children came together over a few years in the largest of these schools. The children were of varied ages (see below for relevant details). In the schools, teachers attempted to teach the children Spanish by using fingerspelling. This is an example of a 'contrived sign system[s] ... based on spoken languages' mentioned by Sandler and Lillo-Martin in the quotation above. The Deaf children communicated spontaneously with one another using **homesigns**, and soon developed a kind of pidgin (which KSC call LSN – see below). This in turn developed into what KSC call ISN, a fully-fledged language, arguably a creolized version of LSN, used to begin with only by children under seven years old. KSC argue that ISN is a new language that was created by the under-sevens at this school, beginning around 1980.

KSC distinguish four types of signing system used in the Deaf community at the school in Managua. First, homesigns, or *mimicas* in Spanish and for the signers themselves. These are ad-hoc signs invented by isolated individuals: 'idiosyncratic gestural systems ... used by isolated deaf

individuals' (180). KSC say '[e]ach isolate's homesign system is unique, idiosyncratic, variable even within the individual, and lacking most characteristics, particularly syntactic, of what we would recognize as a full-fledged human language' (179–80). Homesigns are still in use among the Deaf in Nicaragua, although nowadays they are used only by older signers who made contact with the Deaf community after the critical period for language acquisition was past, and hence they are unable to learn sign language.¹⁷

Most important for present purposes, KSC distinguish *Lenguaje de Señas Nicaragüense* (LSN) from *Idioma de Señas Nicaragüense* (ISN).¹⁸ LSN is 'a highly variable and ever-changing form of communication that developed from the point when the[se] homesigners came together' (180), while ISN is 'a coexisting . . . fully articulated signed language form that is in use only by individuals who entered the schools at ages well below the end of what would count as their critical period for language acquisition' (180). We will look in more detail at some of the structural differences between LSN and ISN below, as well as the age profiles of ISN users.

Finally, KSC distinguish a fourth variety: *El Pidgin de Señas Nicaragüense* (PSN). This is 'a pidgin used between hearing individuals and deaf signers' (181) which 'characteristically involves the interspersing, sometimes overlapping, of Spanish and Nicaraguan signs' (184–5). Interestingly, Deaf signers using PSN consider themselves to be speaking Spanish, while Spanish speakers consider themselves to be signing, but their interlocutors systematically think the opposite. I will have very little to say about PSN here, as it clearly appears to be a pidgin or contact vernacular. Instead, I will concentrate on LSN and ISN, as it is here that the evidence for language creation lies.

KSC show that there are a number of important structural differences between LSN and ISN. First, there is a general difference in the nature of the signing gestures and the signing space (the area in front of the signer's upper body in which signs are made). In LSN 'signs are large and tend to be symmetrical', while in ISN there is a 'smaller signing space . . . The use of the two hands is more asymmetric' (183–4). There are also quantifiable

¹⁷ Such individuals are referred to as NO-SABES ('know nothings') in the sign language. (Here and throughout I follow convention in transcribing signs in capitals).

¹⁸ Here KSC exploit the distinction made in Spanish between *lenguaje* ('language in general') and *idioma*, which designates languages as sociocultural entities.

differences in fluency and rate of information conveyed. (LSN signers averaged 24 events/minute in the study in Senghas *et al.* (1994), while ISN signers averaged 46, $p < .05$.)

Second, LSN and ISN differ in the use of non-manual grammatical markers. This is a fairly common feature of sign languages, originally identified in American Sign Language (ASL) by Liddell (1980); see Sandler and Lillo-Martin (2001: 537–9). It involves a facial expression or head gesture accompanying a manual sign or sequence of signs. For example, in ASL relative clauses are marked by ‘raised brows, a backward head tilt and a tensed upper lip’ (Sandler and Lillo-Martin 2001: 537), and wh-questions by a furrowed brow. The relative-marking extends over the entire relative clause constituent. LSN, unlike *mimicas*, makes use of facial expressions as grammatical markers, but these either occur alone or associated with at most one manual sign. ISN, on the other hand, systematically uses non-manual markers for topics, wh-questions, and yes/no questions. These markers can ‘spread’ over several signs, although how they are bounded is not clear.

Third, LSN makes use of various mouth gestures and vocalizations somewhat sporadically to indicate aspectual categories (intensity, iterativity, and completion). In ISN, on the other hand, these mouth gestures are accompanied by modulations to hand movements and in many cases replaced by them. Iterativity, for example, is marked by verb-reduplication.

Fourth, although ‘in comparing LSN to ISN we don’t see an across-the-board shift from uninflecting to inflecting verbs’ (KSC: 191), there are certain differences in the incidence of inflection, in each case showing ISN to be both richer and more systematic. In sign languages, agreement is expressed by establishing referential index points in the signing space (Sandler and Lillo-Martin 2001: 544–5). Agreement with both subject and object is typically found. Verbs of motion also show ‘locative’ agreement, according to which points in the signing space indicate locations. LSN does not have subject- or object-agreement, but there is sporadic agreement ‘with real-world locations or paths that are in the shared signing space of signer and addressee’ (KSC: 190). Sometimes a complex of pointing gestures following the verb can pick out the participants, as in SPEAK#PersonXPersonY (where ‘#’ indicates that the ‘pointing complex’ is enclitic to the verb). ISN, on the other hand, systematically has person agreement fully incorporated into the verb. KSC observe that ‘it is the dropping of encliticized person markers in favor of spatial incorporation

into the path of the verb (LOC(source)-SPEAK-LOC(goal)) that is characteristic of the shift from LSN to ISN' (191).

Fifth, LSN and ISN differ in their use of object **classifiers**. Classifiers are found in many sign languages, and of course in quite a few spoken languages. Object classifiers are 'typically one-handed, bound morphemes that pick out classes of objects on the basis of physical or abstract characteristics' (194), such as small animal, flat object, long thin object, etc. Both LSN and ISN use such classifiers, but ISN uses them more frequently and systematically, and with sensitivity to argument structure (an object classifier is used where there is no agent).

The sixth and most striking difference between LSN and ISN concerns verbs with more than one argument. LSN observes a general constraint that each predicate can have at most one argument. This naturally gives rise to much verb serialization, although KSC argue that the serial constructions are in many cases only apparent in LSN; instead this is simply parataxis. Thus, for 'the woman pushed the man', LSN has the following kinds of option:

- (30) a. WOMAN PUSH MAN GET-PUSHED
b. WOMAN PUSH MAN REACT
c. WOMAN PUSH MAN FALL
d. WOMAN PUSH MAN CRY
(see KSC, (10), p. 217)

As they point out, '[a]n extra verb, . . . , is often added in order to support the second argument' (218). ISN uses serial-verb constructions very productively (and KSC show that these are true serial constructions, not parataxis), and has 'begun to drop the second verb entirely' (219). This gives rise to 'fully grammaticized transitive verbs' (219), something not found in LSN.

In all then, LSN and ISN differ in the expression of topics and wh-markers, aspectual markers, agreement marking, object classifiers, and transitivity. In every case, ISN shows more grammaticalized constructions, and constructions which are typical of sign languages and are also found in spoken languages. KSC's contention that ISN is a genuine natural language, while LSN is a kind of pidgin, seems to be supported by this evidence.

If ISN is a natural language, how did it develop? According to KSC, it developed initially from homesigns: ISN 'evolved from the jumble of

idiosyncratic homesign/gestural systems in use by students who entered the schools in the late 1970s and early 1980s' (180). They say that the 'first generation of LSN signers were homesigners brought into contact with other homesigners' (187). (They date the first generation to the period 1977–80, since the Managua school was opened under the Somoza regime as a private institution.) They follow Thomason and Kaufman (1988) in taking the PLD for first-language acquirers to be supplied primarily by older children rather than by parents. (See the discussion of Weinreich, Labov, and Herzog (1968) in §4.2.5.) The really key point is what the younger children did with PLD from LSN. In developing the grammatical devices of ISN detailed above, these children 'surpassed their models' and thereby created a new language.

This development, which KSC argue to be a case of abrupt creolization, with LSN serving as a pidgin, led to the creation of a new language. Since LSN and ISN then coexisted, the PLD for future children became much richer. This contention is supported by a study of the use of some of the grammatical features described above in relation to the year an individual first entered the community and in relation to the individual's age at first exposure to LSN. Their findings are summarized by the following quotation (197–8):

The number of inflections and amount of agreement per verb was greater overall for signers who entered the community after 1982 and for signers exposed to language at a young or medium age. Young and medium age signers benefitted particularly by a later year of entry ... Older signers showed no effect of year of entry.

Here 'young' is under 6;6, 'medium' 6;6–10;0, and 'older' above 10;0. We can thus think that ISN first emerged around 1982, two years after the first main influx of children, and that younger children were more able to exploit the enriched PLD that resulted from the development of ISN than older ones.

The idea that approximately seven years old is an important cut-off point also emerges from the discussion of the development of double-valence verbs. Comparing an individual who was nine at the year of entry (1977) with one who was four, KSC found that the nine-year-old produced no occurrences of bivalent verbs, while 13 per cent of possible cases were bivalent in the case of the four-year-old. They comment on this as follows: '[i]f the emergence of double-valence verbs is an indicator of the shift from LSN to ISN, then this places the ability to learn ISN as a first language at somewhere below 7 years of age at entry' (221).

Here KSC are arguing for a strong version of the critical-period hypothesis: the idea that language acquisition is only possible during a certain ‘window’ in childhood, after which the ability atrophies. (This idea was first put forward by Lenneberg (1967).) The point is developed further in the following quotation (203):

Young acquirers (<7) were able to use their innate language capacities to ‘make sense of’ or fill in the grammatical holes in the LSN input they were exposed to, whereas slightly older signers (7 < . . . x . . . <16) could reach a certain consensus with one another on general communication strategies but had to learn to make do with less than the optimal innately determined blueprint for language use. The ‘filling in of holes’ by young acquirers yielded a situation in which learners surpassed their models, acquiring and creating a qualitatively different language.

To put it in terms comparable to those adopted in the previous section, we can say that the under-sevens were able to acquire a system on the basis of the highly deficient PLD. This PLD must have been weakly P-ambiguous in many crucial respects. The younger children were able to cope with this (i.e. ‘fill in the gaps’), while older children were not. We can interpret this as meaning that the parameter-setting capacity is only available during these early years. (This is consistent with what we saw in §5.1: perhaps the learning algorithm atrophies, but not UG.) At this stage of development, even highly impoverished PLD such as LSN can serve as the basis for the emergence of a grammatical system. Later, after age seven according to KSC, although LSN can be learned and used as a vehicle for communication, a true grammatical system like ISN (or any other natural language) becomes inaccessible.

As mentioned above, KSC argue that ISN developed from LSN by a process of abrupt creolization. A consequence of this claim is that creolization does not depend on a substratum, since ISN has no substratum:

there were no substrate ‘languages’ to serve as input to LSN. Early LSN was an outgrowth of the conventionalization that occurred among a critical mass of home-signers, none with mastery of a prior signed language. It required the direct contribution of innate language capacities in its creolization to ISN. (KSC, 212)

Thus KSC’s account of the development of ISN favours the ‘universalist’ position regarding creole genesis. It is also worth noting that, whatever the status of spoken creoles such as those discussed in the previous section, the development of ISN is undoubtedly exceptional; if DeGraff (2005) is right that creoles in general are not exceptional, then KSC’s conclusion may not have a great deal of bearing on how we see the development of spoken creoles.

In fact, ISN has developed in the absence of adstrate influence altogether. Spanish is the only candidate superstrate language, and the children in the schools received a minimal education in lipreading, spelling and pronouncing Spanish. (Recall that LSN and ISN both developed outside the classroom; in the schoolyard and on the buses.) KSC comment, however, that ‘students typically left these schools as functional illiterates with minimal lipreading skills’ (205). In the 1980s, there was no contact with any other signed language, and so no possibility of a signed superstrate. And we have seen that there was no substratum. So ISN may be a case of a creole with no adstrate influence at all.

All of this, particularly the fact that the PLD which led to the creation of ISN among the under-sevens must have been very weakly P-ambiguous, should lead us to expect that ISN instantiates many unmarked parameter settings. This claim is, however, very hard to evaluate for two reasons. First, little information is available about a number of structural features. Second, as KSC point out, signed languages may, owing to the different modality of expression, have a different markedness metric from spoken languages. It is certainly true that signed languages share a number of structural properties, but these are quite distinct for the most part from those Bickerton and others have claimed to characterize creoles. Sign languages tend to have agglutinative morphology, subject-, object- and locative agreement (of the type that seems to have emerged in ISN, see above),¹⁹ and serial verbs (Sandler and Lillo-Martin 2001: 556–8). Of these, only the last is shared with spoken creoles, and may, as we suggested in the previous section, represent an unmarked property. Neither agglutinative morphology nor the very rich agreement system are good candidates for unmarked properties, since both are arguably connected to a high incidence of syntactic movement. (We have seen this in earlier chapters in connection with various kinds of ‘rich’ agreement. See especially the discussion of morphological triggers in §3.5; agglutinating languages may involve a great deal of ‘massive’ movement.) KSC suggest that these properties may be connected to the signed modality since neither the articulators (the arms, hands, and upper body) nor the reception modality, the visual system, are as attuned as the oral/aural modality is to rapid linear processing. Because of this, they suggest, ‘signed languages favor packing more information into a single sign’ (214). Kegl (to appear: 25–6) puts this

¹⁹ This may underlie the fact that ISN allows null arguments (KSC, 192).

point by saying that ‘the demands of vision, the brain architecture dedicated to processing its input, and the physical characteristics of face-, torso- and limb-generated articulation override many of the spoken-creole markedness conditions, particularly those favoring isolating morphology’. This is possible, but it represents a very direct link between the computational system and the modality of transmission. It is unclear why abstract syntactic operations such as Move, Merge, and Agree should be affected by these considerations, unless we can construct a plausible account for the influence of the PF interface (which very clearly has different properties in sign languages as compared to spoken languages) on these aspects of the system. So this question remains open, and, as it stands, ISN may represent a challenge to some of the claims we have made in the foregoing chapters regarding markedness.

In conclusion, it should be clear that KSC’s description of ISN and the circumstances surrounding its emergence are of real importance for linguistic theory for several reasons. These concern the nature of creolization, the theory of markedness, the critical period hypothesis, and the poverty of the stimulus in first-language acquisition. Let us consider these one by one.

First, as we have seen, KSC’s account of the emergence of ISN as a case of abrupt creolization from LSN implies that creolization can take place without adstrate influence of any kind. This is important given the debate between the substratist and universalist positions on creolization, and clearly favours the latter view. Indeed, one could see the Nicaraguan situation as a vindication of Bickerton’s Language Bioprogram Hypothesis, except that this raises the difficulty of accounting for the apparently marked values of certain parameters in ISN.

Second, KSC’s work sheds potential light on markedness theory, although exactly what conclusions to draw here is unclear. It is at least possible that the nature of the modality renders rich agglutinative morphology abundantly available in the PLD, and so, although this is technically a marked property, it is always very robustly triggered owing to the nature of the signed modality, perhaps for reasons along the lines of KSC’s speculation quoted above.

Third, KSC’s results seem to strongly confirm Lenneberg’s critical-period hypothesis. (This is also discussed in Lightfoot (2006: 162ff.), along with other evidence for the critical period.) A crucial aspect of the development of ISN was its acquisition by children under seven. Moreover, as we mentioned, older homesigners who made contact with the Deaf

community after their critical period were unable to acquire ISN. If language represents a distinct cognitive module, as argued in Chomsky (1986) and elsewhere, then it is easy to see how it may flourish at a certain stage of life and then atrophy. However, if central features of the language faculty are not ‘domain-specific’, this view becomes more difficult to maintain. It may be, then, that the critical period relates to the parameter-setting device rather than to the language faculty itself. This would be consistent both with the view put forward by Hauser, Chomsky, and Fitch (2002) that the language faculty is not domain-specific and with KSC’s evidence for a critical period for language acquisition (and with the speculation in §5.1 above regarding the relation between the critical-period hypothesis and the Full Transfer Full Access hypothesis for L2A).

Fourth, arguably the most important conclusion to emerge from KSC’s work is that ‘it doesn’t take language to make language’ (206). The innate capacity for language acquisition will create a grammatical system on the basis of even radically impoverished input such as homesigns or LSN. This is strong confirmation of the correctness of the poverty-of-the-stimulus argument: if there is no innate language faculty, why were the under-sevens so crucial to the development of grammatical features in ISN? And why are those grammatical features so similar to those found in other languages, particularly sign languages? Appeal to functional notions such as communicative needs appears to play little role here: LSN is a perfectly adequate communicative system, it seems, and one could argue that the communicative needs of over-sevens are more sophisticated than those of the under-sevens, yet it is the under-sevens who created the more elaborate new language. LSN is a system created by conventionalization of *ad hoc* homesigns, and meets communicative needs, but when small children are exposed to it, they naturally create a true grammatical system out of it. This remarkable fact is easily explained in terms of an innate language faculty combined with the critical-period hypothesis, but not in more empiricist or behaviourist accounts of first-language acquisition. As Kegl (to appear: 47) puts it: ‘language emergence is basically a case of fooling the Language Acquisition Device (LAD) into thinking it is making generalizations on the basis of language data, even when that data is inconsistent, incomplete or totally lacking.’

However, another very important result is clear from KSC’s study. As they point out: ‘[w]e have discovered that the source of language is within us but that the conditions for its emergence depend crucially upon community’ (223). Isolated Deaf individuals were limited to homesigns; many

of them must have remained language-less for life. Once there was a fledgling Deaf community, however, LSN and then ISN quickly emerged. (Recall that the earliest date for the beginnings of this community is 1977, and that the latest date for ISN to be already in existence is 1982.) This confirms the idea that the innate language faculty requires an environmental stimulus; it is incapable of endogenous development. But the stimulus can (and, indeed, must) be impoverished. The language-acquisition device creates grammatical systems by fixing the parameters of variation: ISN is a very clear case of language creation, creolization in general is a further striking case where, owing to social conditions, the PLD was restricted in certain ways, and the ubiquitous nature of syntactic change itself attests to this constant creation and recreation of grammar. Creating grammars is an instinct that small humans can hardly help but put into practice. All they need is normal contact with slightly bigger humans. To quote Kegl again (to appear: 47): '[a]s long as the signal attended to is produced by humans in a communicative context, the LAD can be fooled'. 'Fooling' the LAD, in this sense, amounts to providing it with something it interprets as PLD; then the parameter-setting algorithm and UG can be set to work and a grammatical system will emerge.

Finally, perhaps the most important conclusion of the research on ISN from the perspective of this book is that 'the child language acquisition process impacts the language and leads to a natural process of historical change over time' (Kegl, to appear: 44). Kegl documents how, in particular since the adoption of ISN as the language of the schools in the 1990s led to hearing teachers, native speakers of Spanish, using ISN, there has been evidence of contact effects from Spanish. Kegl comments that 'change is a fact of life in language emergence. Five to ten years can yield a body of data that bears little resemblance to its predecessor' (45). This confirms the old idea that variation and change are inevitable, and the more precise notion that the PLD, the learning algorithm, and the set of available grammars together form a dynamical system. (The propensity for change was present in the initial conditions, given the widely varying nature of homesigns – see Kegl (to appear: 48–51).)

5.5. Conclusion to Chapter 5

This chapter has focused on language contact in relation to syntactic change. What we have seen is that the model of how contact may affect

the PLD, either directly or indirectly, given in (1) is highly relevant for many cases of syntactic change. First, we saw how recent studies of interlanguage confirm the idea that contact situations can significantly perturb the PLD and give us an idea of what imperfect learning might mean. Second, we considered some putative and actual cases of contact and possible substratum effects. Third, we looked at creoles and creolization, concluding, along the general lines of DeGraff (2005), that creoles may not be quite as exceptional in their historical development as has been claimed both by advocates of a language bioprogram such as Bickerton and by advocates of radical relexification. Instead, it may be that, owing to the extreme social conditions under which creoles are formed, we are dealing with a case of radically impoverished PLD due to highly imperfect learning of a second language (i.e. pidginization). Finally, §5.4 considered a truly exceptional situation: that of the Deaf children in Nicaragua, who appear to have spontaneously developed a natural language. Interestingly, here too the notion of impoverished PLD is relevant. For the first group of children the PLD was radically impoverished, in that it consisted purely of homesigns, while for the younger children who came into the schools after 1982, the PLD was more like a pidgin, which they effectively creolized. In both cases, the PLD was replete with weak P-ambiguity, we can assume.

In all of these cases, we are dealing with grammatical expansion: this is clearest in the Nicaraguan case, but also holds of creoles, and possibly of any case where acquirers are exposed to PLD which consists in whole or in part of interlanguage. It may be that these situations simply reveal with greater clarity what happens in all cases of language acquisition and change: that each individual creates a grammar afresh on the basis of fragmentary, impoverished, and noisy experience. In a sense, the truly startling fact is that the Inertia Principle can even be contemplated, since this can be taken to assert, as we have seen, that most of the time language acquisition is convergent. That acquirers can achieve this at all is *the* linguistic phenomenon most in need of explanation, as Chomsky has repeatedly emphasized in his writings.

In all the cases of grammatical expansion, we have weak P-ambiguity in the PLD. However, we saw in §2.1 that abductive reanalysis involves strong P-ambiguity. It may be worth pursuing the question of whether the types of ambiguity naturally favour different types of change. This is a question which arises from what we have seen, but it is not one which I will pursue here.

Further reading

Pidgins and Creoles

Holm (1988) is a very wide-ranging survey of all spoken creoles about which information was available at the time. It is an invaluable starting point for any investigation into the nature and properties of creoles. **Mühlhäusler (1986)** is a useful introduction to pidgins and creoles and the issues surrounding them. **Romaine (1988)** is another very useful general introduction to pidgins and creoles. **Green (1988)** is a survey of Romance-based creoles, with much useful information regarding their syntactic properties. **Schuchardt (1979; 1980)** are translations of the works of Schuchardt originally written in the last decades of the nineteenth century. Schuchardt pioneered the study of pidgins and creoles, and raised many of the questions regarding their origin and nature which are still discussed today. **Thomason and Kaufman (1988)** surveys the field of contact linguistics from a sociolinguistic, historical linguistic, and typological point of view. Their ninety-page discussion of the sociolinguistic situation in Medieval England is extremely useful. **Thomason (2003)** is a general overview of the literature on the relation between language contact and language change. **Bickerton (1981)** is where the Language Bioprogram Hypothesis is put forward in detail for the first time. **Bickerton (1984)** is a summary of the principal ideas; here principles-and-parameters theory is taken into consideration and Bickerton suggests that creoles might correspond to largely unmarked values of parameters. See also **Bickerton (2004)**. **Lefebvre (1998)** is an extensive analysis of Haitian, in which the West African substrate is defended in detail as the explanation for many of the respects in which Haitian differs from (Modern) French. **Lefebvre and Lumsden (1989)** is an early statement of the same idea, and **Lumsden (1999)** further develops some of these ideas. **Koopman (1984; 1986)** are early generative studies using the government-and-binding model both of the Kwa languages and of Haitian. Koopman argues that the overall syntactic similarities between these languages justify the postulation of a West African substrate for Haitian. **Mufwene (1986)** argues that, in order to fully understand the nature of creoles, a combination of the universalist (i.e. Bickertonian language bioprogram) and the substratist positions is required. **Mufwene (2001)** sustains this argument. **DeGraff (1994)** and **DeGraff and**

Dejean (1994) both look at V-to-T movement in Haitian and other French-based creoles. **DeGraff (2003; 2004; 2005; to appear)** are more recent papers in which DeGraff makes the case against what he calls ‘creole exceptionalism’: the idea that creoles are exceptional either by virtue of revealing the language bioprogram or UG in any special way or by virtue of having undergone especially radical relexification. **Rizzi (1999)** is one of the Epilogue chapters to DeGraff (1999) (see the further reading to Chapter 3), in which a number of issues common to creolization and first-language acquisition are discussed, notably the question of what he calls ‘grammatical expansion’.

The Celtic languages and English

Cottell (2002) looks at clefting constructions in Hiberno-English, and suggests that some of their properties may be attributable to substratum effects from Irish. **Thomas (1997)** is a useful survey of Welsh English, with some very interesting discussion of possible substratum effects, some of which are taken up in §5.2.3. **McCloskey (1992)** studies a variety of Hiberno-English in which subject-auxiliary inversion (or T-to-C movement) is allowed more freely than in Standard English, but still not in embedded semi-questions, in the sense introduced in §5.2.3. **Henry (1995)** analyses a number of constructions of Belfast English which are not found in Standard English in some detail, although she takes no position on whether they can be attributed to substrate influence. **Tallerman (1996)** is a detailed study of the different kinds of fronting constructions (topicalization and focalization) available in Modern Welsh. Tallerman argues that at least two CPs are required in order to capture the facts. **Roberts (2005)** is a detailed study of Modern Welsh, concentrating on clause structure and the derivation of VSO order. **Rouveret (1994)** is another detailed study of Welsh, covering slightly more empirical ground than Roberts. **Ó Murchú (1993)** is a general description of Modern Irish, with some useful detail regarding the sociolinguistic situation. **Owen Jones (1993)** is a similar description of Modern Welsh.

Other work on the history of English

Jespersen (1938) is a fairly elementary general introduction to the history of English, which contains much useful discussion of changes in vocabulary.

Kastovsky (1992), a contribution to the *Cambridge History of the English Language*, looks in some detail at the sociolinguistic situation in England at the time of the Danelaw. **Kroch, Taylor, and Ringe (2000)** argue for a contact-based account of certain changes in ME syntax, involving dialect contact between Northern and Southern dialects of ME. **Trips (2002)** is an in-depth study of the *Ormulum*, a text probably composed in Bourne, Lincolnshire, around 1180, i.e. in the Danelaw, and clearly subject to Scandinavian influence, at least in vocabulary. Trips shows that the text has a number of syntactic features that are likely to be of Scandinavian origin, and suggests that these support the hypothesis that the word-order change from OV to VO in English may have been due to Scandinavian influence.

Nicaraguan and other sign languages

Kegl, Senghas, and Coppola (1999) is the main article on Nicaraguan Sign Language. Here the claim that ISN emerged among the Deaf children in Managua is substantiated in detail. **Senghas (1995a, b)** are in-depth studies of different aspects of LSN and ISN, in which the differences between the two varieties are brought out clearly. Much of the data used to support the arguments in Kegl, Senghas, and Coppola (1999) originated here. **Kegl (to appear)** summarizes and elaborates some of the results of the earlier work on ISN, as well as providing a brief history of ASL. **Liddell (1980)** was a pioneering study of American Sign Language, in which it was argued that this language shared many of the important structural features of spoken languages. **Sandler and Lillo-Martin (2001)** is a state-of-the-art survey of what is known about the structure of sign languages, focusing primarily but not exclusively in American Sign Language.

The biological foundations of language

Lenneberg (1967) is a classic statement of the biological foundations of language. Here the critical period hypothesis – the idea that the ability to acquire a language under naturalistic conditions is lost at or sometime before puberty – was first proposed and argued for.

Epilogue

Although I said in the Introduction that my goal was to extend, not to defend, Chomsky's thinking on language, what really emerges from the foregoing is the importance of the poverty of the stimulus to language acquisition. Because grammars are recreated by each cohort of acquirers within the tight constraints imposed by UG, the variation and change that is so prevalent in language arises. All that seems to be required for this is for certain parts of the system to be indeterminate (for example, the nature of certain features associated with certain heads) and, given the indeterminacy of the PLD, different adult systems will emerge. These systems stabilize as such after the critical period, and typically become associated with social and cultural value (in an ultimately quite arbitrary way, as far as the system itself is concerned). The ongoing, inevitable propensity for variation, determined by the underspecified parts of the formal system, leads to the creation of new systems which fit into the social value system in different ways. Thus a parameterized UG, allowing random variation in a few small areas, gives rise to the phenomena of variation – both sociolinguistic and cross-linguistic – and change. And children have the ability to acquire these systems along with the variation and, under conditions whose precise nature I have tried in the foregoing to delineate, although so much remains obscure, they have the capacity to subtly modify the system and so innovate.

The study of historical syntax can, in these terms, find its natural place in the cognitive sciences. What will hopefully develop is a greater understanding of each of the three elements which, according to Niyogi (2004), contribute to the dynamical system that is a language being spoken by a population. We need to better understand the nature of language learning and acquisition through empirical work on L1A and L2A and theoretical work on learnability; we need to better understand the relation between language variation and populations through empirical work in sociolinguistics and

theoretical work on the computational modelling of population dynamics. Finally, we need to better understand the properties of the set of grammars through empirical work in language typology and theoretical work in the theory of UG. Principles and parameters theory is obviously central to this last enterprise, itself crucial to a full understanding of language change. As Lightfoot (2006: 166) says, if ‘we are generalists in this way, then linguists can attain a level of explanation quite unlike what one finds in historical studies in other domains, such as the theory of biological species or political systems.’

If this book is to have a general conclusion, then, it may be this: that the variation and change that is prevalent in language finds its natural explanation within a Chomskyan linguistic paradigm. I have certainly not been able to prove this thesis in the foregoing, but I hope to have done enough to show that the idea is worth investigating, and that the existence of variation and change in language does not in any way argue against the generative approach to explaining language. Quite the contrary, in fact.

Glossary

Abductive change: change caused by the fact that learners only have access to the output of a generative grammar (*q.v.*) and to Universal Grammar (*q.v.*) with no direct access to the grammar itself. The combination of primary linguistic data (*q.v.*) and Universal Grammar may lead the learner to abduce a system which is distinct from that underlying the primary linguistic data by reanalysis (*q.v.*). Since Andersen (1973), abduction has been recognized as a potentially important mechanism of language change.

Agree: in the Minimalist Program (*q.v.*), a matching relation holding between formal features in a particular syntactic domain. One term, α , Agrees with another term, β , iff α asymmetrically c-commands β , α and β are non-distinct in formal features and they are in a specific local domain. See §1.4.1 for the definition of asymmetric c-command ((90)) and the local domain ((89iii)).

Algorithm: ‘a clerical or effective procedure that can be applied to any of a class of certain symbolic inputs and that will in a finite time and number of steps eventuate in a result in a corresponding symbolic output’ (Audi (1999: 21–2)). Turing machines are abstract machines capable in principle of carrying out any algorithm. Generative grammars (*q.v.*), as a subset of the set of Turing machines, are algorithms.

Argument structure: the number and nature of the participants implied in the eventuality described by a lexical word: *donate* implies three, *transplant* implies two, *die* implies one. Verbs are usually thought to have a richer argument structure than other syntactic categories.

Binding theory: a ‘module’ of government-binding theory which deals with the nature of and constraints on anaphoric relations involving DPs of various kinds: reflexives (*himself*, etc.), reciprocals (*each other*, etc.), pronouns (*me*, *you*, *his*, etc.), and ‘referring expressions’ or non-pronominal DPs (*John*, *the man*, *Bill’s liver*, etc.). Minimalist approaches tend to view the principles of the earlier binding theory as high-level descriptive generalizations in need of explanation.

Classifiers: (more precisely, sortal numeral classifiers, Gil (2005: 226)). DP-internal morphemes, common in East Asian languages but found in

other areas (for example, in Amazonian languages), which co-occur with lexical nouns in construction with a numeral and vary according to some generic feature of the noun's meaning; often required in plural or quantified expressions where nouns do not exhibit a mass-count distinction of the type familiar from English. The nearest English equivalents are examples such as *three head of cattle*. ASL and other sign languages, including ISN, have rich classifier systems.

Clitic: a morpheme phonologically, but perhaps not syntactically or morphologically, dependent on another element, known as the host. The English contracted auxiliaries, for example, *'ll* in *they'll transplant it*, are enclitic on the subject or a fronted *wh*-expression. 'Special' clitics occupy designated positions, for example, the pronominal clitics of the Romance languages which are almost always dependent on a verb, or the second-position clitics found in some South Slavonic languages and elsewhere.

Code-switching: (also known as 'code-mixing'), the phenomenon of using more than one language in the course of a single dialogue, monologue, or sentence. It is clear that speakers can switch from one grammatical system to another in mid-sentence, although it is unclear whether there are structural constraints on where the switch can be made.

Complementizer: a subset of what are known as subordinating conjunctions in traditional grammar. Complementizers are typically sensitive to clause type, for example, whether a clause is interrogative, declarative, exclamative, etc. They usually appear in complement clauses, where they are selected by the superordinate predicate, and in adjunct clauses of various kinds, although in some languages they systematically appear in main clauses. In current syntactic theory, the complementizer position is the C head of CP, which takes TP as its structural complement.

Creoles: languages historically derived from pidgins (*q.v.*), and as such perhaps having a 'break in transmission' in their history. For this reason, creoles are of particular interest in historical linguistics.

Creolization: the process of formation of a new creole (*q.v.*), often, perhaps always, from a pidgin (*q.v.*). The extent to which creolization reveals or obscures the usual processes of first- or second-language acquisition is much debated.

Critical-period hypothesis: the idea, originally put forward by Lenneberg (1967), that 'there is a time period which is optimal for language acquisition, with a maturational decline with increasing age' (White 2003: 245). L2 acquisition takes place after the critical period, it is often assumed. Niyogi

(2004) shows that the critical-period hypothesis may be sufficient to guarantee variation in a speech community after a single generation (see §4.2.3). The invention of ISN by Deaf children in Nicaragua appears to support the critical-period hypothesis (see §5.4). This hypothesis comports well with the innateness hypothesis (*q.v.*), although neither entails the other.

Cross-categorical harmony: the observation, formulated and labelled as such by J. Hawkins (1983), that the Greenbergian dyads (VO/OV, Prepositions vs. Postpositions, NRel vs. RelN order, etc.) tend to pattern together. The theoretical interpretation of this tendency remains unclear.

Cue: in syntax ‘a piece of structure . . . which is derived from the input, but . . . is not a sentence’ (Lightfoot 2006: 78); Dresher (1999) sees cues as triggering parameters, part of the statement of phonological parameters (for example, Quantity (In)sensitivity, as in (31) in §3.3.2).

Diary drop: the appearance of null subjects in otherwise non-null-subject languages in colloquial written registers such as diaries and personal correspondence, and perhaps in informal colloquial speech. Null subjects in these contexts appear to be subject to a particular set of constraints, rather different from those affecting null subjects in canonical null-subject languages like Italian.

Diglossia: the sociolinguistic phenomenon whereby two distinct languages or varieties are used by members of a single speech community for clearly circumscribed purposes, usually one ‘high’, or relatively formal and impersonal, and the other ‘low’, or relatively informal or intimate.

Discrete: symbols or other entities are discrete if they are clearly distinguishable from one another. In standard conceptions of phonology, phonemes are discrete entities; a glance at a spectrogram, however, shows us that the physical sounds of speech are not discrete. The natural numbers are discrete, and the number system has the property of discrete infinity, in that the set of these numbers is unbounded. Chomsky has frequently suggested that humans uniquely have the cognitive capacity to apprehend discrete infinity, both in our numerical and in our syntactic competence. (See for example (2002: 45–6).)

Dynamical systems: ‘any evolving ensemble where variation of a parameter-setting produces a change of state’ (Lass 1997: 293). (Here the term ‘parameter’ is used in its usual mathematical sense, not in the sense specific to linguistic theory as in this book; on the connection between the two, see Baker (2001:55)). A dynamical system ‘can be mapped as trajectory in a multidimensional space (‘phase-space’) where each point in the space

represents a possible system-state' (Lass 1997: 293). Niyogi (2004) shows at length how the combination of a set of generative grammars (*q.v.*), a set of learning algorithms (*q.v.*), and a random distribution of primary linguistic data (*q.v.*) is a dynamical system. Many natural phenomena can be modelled as dynamical systems: see in particular Niyogi's (2004: 454–57) comparison of language acquisition under certain conditions with the behaviour of particles in magnetic fields.

Factive: a factive predicate presupposes the truth of its sentential complement, as in *I (don't) regret that he had a liver transplant*. Verbs such as *think* are non-factive, on the other hand: compare *I (don't) think that he had a liver transplant*, where there is no commitment to the truth or falsity of the proposition expressed by the subordinate clause.

Formal: this adjective is often taken to mean something similar to algorithmic (*q.v.*), in that it designates an approach which is precise and rigorous, perhaps involving the use of mathematical notation, thus generative grammar (*q.v.*) is a formal theory of grammar; more substantively, it refers to a grammatical description which is concerned with the linguistic form rather than linguistic function. In this sense, too, generative grammar is a formal theory.

Formal universals of language: the formal concepts, in the second sense of formal (*q.v.*), which make up Universal Grammar (*q.v.*), for example, Merge (*q.v.*), Agree (*q.v.*), Move (*q.v.*), etc.

Generative grammar: the theory of language put forward by Chomsky (See the Readings to the Introduction, under the heading 'Chomsky's work and introductions to it'.) Generative grammars recursively enumerate the set of well-formed expressions in a language, along with their structural descriptions. The well-formed sentences of the language are akin to the theorems of a deductive system, and the rule systems that specify them are akin to rules of inference. The theory of generative grammar aims to provide an account of the human language faculty (*q.v.*) by specifying the class of possible grammars of individual languages that Universal Grammar (*q.v.*) allows. This faculty underlies the human capacity to acquire language and is hypothesized to be a genetically-inherited species characteristic of humans, even if aspects of it may have homologues in aspects of the cognitive systems of other animals.

Grammaticalization cycles: cases where different lexical elements have become grammaticalized in a given function at successive periods, with one consistently in a less grammaticalized stage than the other. Jespersen's

Cycle of Negation is perhaps the most famous example of a grammaticalization cycle.

Grammatical functions: (also known as grammatical relations) notions such as subject-of, direct-object-of, predicate, etc., which played a central role in traditional grammar. In generative grammar (*q.v.*), grammatical functions are usually defined in terms of phrase-structure configurations. This is not the case in certain variants, such as Relational Grammar, Arc-Pair Grammar, and Lexical-Functional Grammar. To be kept distinct from thematic roles (*q.v.*).

Homesigns: ad-hoc signs invented by isolated Deaf individuals, such as congenitally Deaf children born to hearing parents who are unable to acquire spoken language since they cannot access the primary linguistic data (*q.v.*) ('idiosyncratic gestural systems . . . used by isolated deaf individuals' (180)). Kegl, Senghas, and Coppola (1999) argue that homesign systems are 'lacking most characteristics, particularly syntactic, of what we would recognize as a full-fledged human language' (179–80). Goldin-Meadow (2005) takes up this point, and suggests that homesigns have some properties we might expect from a system determined by Universal Grammar (*q.v.*) (218–19).

Implicational universals: statements of the general form 'if a language *L* has property *p*, then it has property *q*'. Originated in Greenberg (1963) (many of whose implicational statements were probabilistic, containing the proviso 'with greater than chance frequency'), implicational universals have been a central topic in language typology (*q.v.*). Their importance is that, by excluding (or rendering unlikely) one of the four logically possible combinations of properties, they tell us that grammatical variation across languages is not random.

Inertia Principle: the idea that, unless some force acts upon a grammatical system, it will not change. It may be interpreted as asserting that, in general, language acquisition converges successfully on the target system. In §4.3 we entertained the possibility that inertia might entail a 'path of drift' through the state-space defined by the set of UG parameters (taking parametric systems to be dynamical systems (*q.v.*)) towards an attractor.

Innateness hypothesis: the claim that language acquirers bring some predisposition to the acquisition task. Chomsky claims that this is Universal Grammar (*q.v.*), and defends this with the argument from the poverty of the stimulus (*q.v.*). Accounts of language acquisition which assume absolutely no innate predisposition to language are rare and unsuccessful.

Interlanguage: the version of language L produced by learners of L for whom L is not their native language (first language, or L1). Questions arise as to whether interlanguage can or must be based on Universal Grammar (*q.v.*), the extent and nature of L1 interference, and the relation between interlanguage and (possibly intermediate) stages of L1 acquisition. Interlanguage may be important for syntactic change since contact situations may give rise to interlanguage as primary linguistic data (*q.v.*) for L1 acquisition, with possibly important consequences for the final state of that acquisition process.

Islands: syntactic domains out of which movement, in particular wh-movement, is not possible. See Box 1.6.

Isolating, agglutinating, and inflectional languages: the morphological typology put forward in the early nineteenth century by Schlegel and modified later in that century by Schleicher. Isolating languages make little or no use of inflection: the standard example is Vietnamese. Agglutinating languages make use of inflections fairly transparently attached to a root in a (near) one form–one meaning relation; the standard example is Turkish. Inflectional languages (also known as fusional languages) attach inflections to a root in more opaque fashion, with no one-to-one form–meaning relation; Latin is the standard example. For more discussion and illustration, see Comrie (1989:42ff.).

Labelled bracketing: one way of presenting the structural description of a sentence in generative grammar (*q.v.*), with brackets showing the boundaries of constituents and subscripted labels indicating the category of each constituent, as in [_{VP}[_{VP}go] [_{PP}[_{PTO}] [_{DP} hospital]]]. Equivalent to a tree diagram (*q.v.*).

Language faculty: whatever cognitive structure underlies our ability to acquire our native language, to store the knowledge so acquired in the mind/brain and to put it to use in production and comprehension. In generative grammar (*q.v.*), Universal Grammar (*q.v.*) is the theory of the human language faculty.

Language typology: ‘the scientific study of variation and the limits to variation in the structure of languages’ (Haspelmath 1998: 8). Language typology can clearly inform our picture of Universal Grammar (*q.v.*), particularly given the principles-and-parameters perspective.

Learnability: the property of a grammar which makes it attainable by a learning algorithm (*q.v.*) on the basis of plausible primary linguistic data (*q.v.*). The basic problem is that of ‘identifying an unknown set on the basis

of example sentences' (Niyogi 2004: 54). It is usually thought that the learner must have some kind of disposition to learn a particular kind of grammar, favouring the innateness hypothesis (*q.v.*) in some form, otherwise the poverty of the stimulus (*q.v.*) makes the learning problem insoluble (see Niyogi (2004: 16)).

Learning algorithm: 'an effective procedure [or algorithm (*q.v.*) – IGR] allowing the learning child to construct hypotheses about the identity of the target language on the basis of examples it has received' (Niyogi 2004: 57). Given the assumptions made in this book, the learning algorithm takes primary linguistic data as input and yields a grammar, an instantiation of Universal Grammar (*q.v.*) with all parameters set to determinate values, as output. Children in the critical period (*q.v.*) have access to such an algorithm.

Logical negation: in standard propositional and predicate logic, the constant which, applied to a proposition, changes its truth value from 1 to 0 or from 0 to 1. Usually written \sim or \neg . In natural languages, clausal or sentential negation is thought to approximate logical negation in its usual use.

Logical problem of language change: in the context of the idea that language change arises through the language-acquisition process, the problem of why acquirers would converge on a system different from that which produces the primary linguistic data they are exposed to: if that system generates the data, how are acquirers led to postulate a distinct system? See Clark and Roberts (1993: 300).

Markedness: in essence, the intuition that a binary opposition is asymmetric, in that one of the terms is in some sense more complex than the other. The more complex term is the marked term of the opposition, the simpler one is the default. A symmetric opposition in which both terms are equal in markedness is known as an equipollent opposition. It has frequently been suggested that parameters may have marked and unmarked values, an idea explored at length in §3.4 and §3.5.

Markedness reversal: a change such that the formerly marked term of an opposition becomes unmarked and vice versa. An interesting example, from phonological change in the history of English, is discussed in Kiparsky (2003).

Merge: the basic structure-building operation in minimalist syntax. Merge combines two terms α and β , to form a third term γ with label δ . Formally, γ consists of the set consisting of the label and the set $\{\alpha, \beta\}$,

i.e. $\{\delta\{\alpha, \beta\}\}$. Both α and β may have internal structure. In other words, Merge is recursive (*q.v.*), in the sense that it applies to its own output.

Minimalist Program: the current version of generative grammar (*q.v.*), originated in Chomsky (1993) (reprinted as Chapter 3 of Chomsky (1995)). The leading idea is that syntactic operations should be reduced to ‘(virtual) conceptual necessity’, meaning that as few specifically syntactic mechanisms should be postulated as possible (methodological minimalism, equivalent to Occam’s Razor), and that the fundamental properties of the model of syntax should follow from principles of optimal design, etc. (substantive minimalism). Chomsky (2002, 2004, 2005a,b) are recent articulations of the point of view.

Move: the mechanism by which elements are displaced from one position to another in a syntactic derivation or representation. Move consists of copying an element in a new position (also called ‘internal Merge’) and deleting the original copy (at PF). See Box 1.1 for more details.

Negative evidence: putative aspects of primary linguistic data (*q.v.*) which inform acquirers of what is not possible. Generally thought to play no role in L1 acquisition.

Object control: the interpretative relation between the direct object of a superordinate clause and the understood subject of an immediately subjacent infinitive, as in [*John persuaded Mary [(subject) to get a liver transplant]*].

Observational, descriptive, and explanatory adequacy: the three levels of adequacy for generative grammars (*q.v.*) originally defined in Chomsky (1964). An observationally adequate grammar correctly distinguishes the well-formed strings from the ill-formed ones, in accord with native-speakers’ intuitions about grammaticality, by generating all and only the well-formed ones. A descriptively adequate grammar does the same as an observationally adequate one, but specifies the correct structural descriptions of the strings at the same time. An explanatorily adequate theory explains how a descriptively adequate grammar can be acquired; in other words it derives the descriptively adequate grammar directly from Universal Grammar (*q.v.*). Chomsky (2004) suggests that the Minimalist Program may take us beyond explanatory adequacy. Longobardi (2003) suggests that historical linguistics raises its own adequacy criteria (see §4.4.6).

OV language: a language in which, in canonical order, the object precedes the verb. About half of the world’s languages are OV, including major languages like Japanese, Korean, and Turkish (see §1.6.1). OV order is

associated with a number of implicational universals (*q.v.*), giving rise to cross-categorical harmony (*q.v.*).

Parameter of Universal Grammar: a choice-point or open option in Universal Grammar (*q.v.*). Parameters are the principal construct in the analysis of cross-linguistic variation, both synchronic and diachronic, in both minimalism and government-binding theory, and form a central theme of this book.

Pidgin: a communicative system which may form the basis of creolization (*q.v.*). Pidgins may not be natural languages, and are usually thought not to have native speakers.

Pied-piping: the phenomenon whereby a category larger than one would expect undergoes Move. The metaphor is that the element for which movement is motivated acts like the Pied Piper of Hamlyn, in taking the extra things away from where they belong. The classic example of pied-piping appears when prepositions move along with their *wh*-complements, as in [_{PP} *To whom*] *do you wish to donate that organ?* (compare *Who do you wish to donate that organ to?* where there is no pied-piping).

Polarity items: words or phrases which depend on a superordinate negative or positive element for well-formedness. An idiomatic phrasal negative polarity item in English is *lift a finger* as in *John didn't lift a finger to help me* (compare *John lifted a finger to help me*, which only has the literal meaning).

Poverty of the stimulus: the most important argument for the innateness hypothesis (*q.v.*). Often misunderstood as asserting that certain types of strings cannot be present in the primary linguistic data (*q.v.*), it states that primary linguistic data inevitably underdetermines the abstract mental grammar that can be constructed from it if the learning device has no built-in predispositions. Given a *tabula rasa*, the hypothesis space is too large for anything like a natural-language grammar to have any chance of being constructed on the basis of experience of the primary linguistic data alone. See Chapter 1 for discussion.

Presupposed: *p* presupposes *q* if the truth of either *p* or its negation guarantees the truth of *q*. For example, both *the King of France exists* is true if either *the King of France is bald* or *the King of France is not bald* is true. Presupposition differs from entailment in that if *p* entails *q*, not-*p* does not have to entail *q*. The truth of the proposition expressed by the clausal complement of a factive (*q.v.*) predicate is presupposed.

Primary linguistic data (PLD): tokens of linguistic behaviour on the basis of which a learning algorithm (*q.v.*) can construct a grammar.

Natural-language PLD is thought to consist only of positive evidence (*q.v.*). Subtle changes to PLD may give rise to language change.

Probably Approximately Correct (PAC) algorithms: a class of learning algorithms (*q.v.*) which identify a probabilistic characterization of learnability. These can be used to characterize ‘the probability with which a typical child might acquire the target grammar after its critical linguistic experience during the learning phase’ (Niyogi 2004: 85). As such, PAC algorithms can play a role in modelling change driven by acquisition.

Quantificational force: the intrinsic content of a quantifier like \forall in predicate logic or *every* in English. According to standard formal semantic theory, quantifiers denote relations among sets: *every* for example denotes the subset relation in that if *every man is a fool* is true then the set of men is a subset of the set of fools. The quantificational force derives from the relation between the two sets that constitute the arguments of the quantifier.

Quantified expressions: in predicate logic, formulae containing at least one of the quantifiers and a variable bound by it, for example, $\forall x[F(x)]$. In natural languages, typically (but by no means exclusively) sentences whose subject DPs contain what are traditionally called indefinite pronouns, adjectives, or articles of various kinds, for example, *Everything is in a state of flux*.

Reanalysis: a central concept in diachronic linguistics, the notion that acquirers may assign a structural description to a string which differs from the one in the grammars of those who produce the string in the PLD (*q.v.*). A pair of grammars G_1 and G_2 such that G_2 contains reanalysed strings of G_1 (say G_2 assigns the structure $[a[ab]]$ to the string *aab* while G_1 assigns it $[aa[b]]$) may have the same weak generative capacity, in that they generate the same set of strings, but they differ in strong generative capacity in assigning different structural descriptions to these strings. In the case of syntactic change, then, individuals may appear to be speaking the same language, in the sense that they are producing the same strings, but actually have different grammars in terms of strong generative capacity.

Recursive: recursion is the property of a formal system which allows it to apply to its own output. The rule systems of generative grammar (*q.v.*), including Merge (*q.v.*), are recursive. This is important as it allows us to capture the fact that natural syntax is unbounded (every language has an infinite number of grammatical sentences) with a finite rule system. Recursion also underlies discrete (*q.v.*) infinity.

Relexification: the process whereby the syntax of a language remains constant while the vocabulary is replaced. This is thought to be one of the mechanisms of creolization (*q.v.*) by substratists (see §5.3.3).

Root infinitives/optional infinitives: a phenomenon in L1 acquisition involving the production of infinitive clauses as declarative main clauses. Thought to be restricted to non-null-subject languages, and characteristic of the second and third years of life. See Guasti (2002) and the discussion and references in §3.1.

Sign language: language produced through the visual/gestural modality rather than the oral/aural one, typically but not exclusively by hearing-impaired individuals or communities. It is now accepted that sign languages are natural languages, although there are artificial sign languages which are not (for example, fingerspelling). Some gestural systems are probably not natural languages, for example, homesigns (*q.v.*), although they can be the precursors to natural sign languages under certain conditions. See Goldin-Meadow (2005).

Subjacency condition: a principle of grammar, thought to be part of Universal Grammar (*q.v.*), which can describe many of the facts related to islands (*q.v.*). Some details are given in Box 1.6.

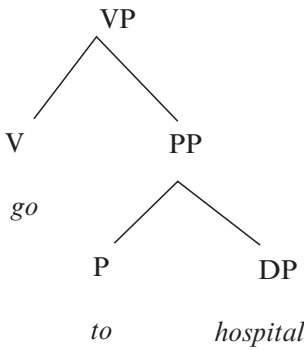
Subject control: the interpretative relation between the subject of a superordinate clause and the understood subject of an immediately subjacent infinitive, as in [*John* tried [*(subject)* to get a liver transplant]].

Substantive universals of language: the substantive concepts, probably encoded as features, which make up Universal Grammar (*q.v.*), for example, negation, tense, and quantification.

Stratum: the typical case involves a language shift by a given population from an ancestral language to a new language, such that the new language, although not descended from the ancestral language, is influenced by it often in rather subtle ways. Gaulish forms a substrate for French, Brythonic Celtic for English, perhaps Basque for Spanish, etc. In the case of creoles (*q.v.*), the ancestral language of a displaced population who were later exposed to a pidgin (*q.v.*) which they creolized (*q.v.*).

Thematic roles: semantic relations such as Agent, Patient, and Recipient, associated with the arguments in the argument structure (*q.v.*) of a predicate. To be distinguished from grammatical functions (*q.v.*).

Tree diagram: one way of presenting the structural description of a sentence in generative grammar (*q.v.*), with superordinate nodes showing constituents and appended labels indicating the category of each constituent, as in:



Equivalent to a labelled bracketing (*q.v.*).

Unaccusative: the type of intransitive verb whose single argument is merged as a direct object, for example, *die*. Unaccusatives do not assign an Agent thematic role (*q.v.*) to their argument.

Unbounded dependencies: syntactic relations, including some types of Move, which appear to hold over an arbitrarily long stretch of intervening material. The existence of islands (*q.v.*) shows that unbounded dependencies are tightly constrained. The subadjacency condition (*q.v.*), in its standard form, effectively claims they are illusory. See Box 1.6.

Unergative: the type of intransitive verb whose single argument is merged as a subject, for example, *shout*. Unergatives typically assign an Agent thematic role (*q.v.*) to their argument.

Uniformitarian hypothesis: the idea that the languages of the past are not essentially different from the languages of the present. In terms of principles and parameters, this can be thought of as the claim that the languages of the past reflect the same Universal Grammar (*q.v.*) as those of the present, but perhaps with differing parameter settings. An essential idea for historical linguistics, it was not systematically applied until the nineteenth century.

Universal Grammar: the theory of the human language faculty (*q.v.*). Universal Grammar is usually thought to consist of invariant elements of various kinds, associated with a restricted domain of variation described by parameters (*q.v.*).

VO language: a language in which, in canonical order, the object follows the verb. Just under half of the world's languages are VO, including English and the Romance languages (see §1.6.1). VO order is associated with a number of implicational universals (*q.v.*), giving rise to cross-categorial harmony (*q.v.*).

References

- AARSLEFF, H. (1970). 'The History of Linguistics and Professor Chomsky', *Language* 46: 570–85. Reprinted in H. Aarsleff (1982), *From Locke to Saussure*. London: Athlone.
- ABOH E. (1999). *From the Syntax of Gungbe to the Grammar of Gbe*. PhD Dissertation: University of Geneva.
- (2005). 'Object Shift, Verb Movement, and Verb Reduplication', in G. Cinque and R. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press, pp. 138–77.
- ADAMS, M. (1987a). 'From Old French to the Theory of Pro-Drop'. *Natural Language and Linguistic Theory* 5: 1–32.
- (1987b). *Old French, Null Subjects and Verb-Second Phenomena*. PhD Dissertation: UCLA.
- (1988a). 'Embedded *Pro*', in J. Blevins and J. Carter (eds.), *Proceedings of NELS 18*, GLSA, University of Massachusetts, Amherst, pp. 1–21.
- (1988b). 'Les effets V2 en ancien et en moyen français', in P. Hirschbuhler and A. Rochette (eds.), *Aspects de la syntaxe historique du français*, *Revue québécoise de linguistique théorique et appliquée* 7: 13–40.
- ADGER, D. (2003). *Core Syntax*. Oxford: Oxford University Press.
- D'ALESSANDRO, R., and I. ROBERTS. (2006). 'Split Ergativity in Abruzzese and the Null-subject Parameter'. Paper given at the 34th Romance Linguistics Seminar, University of Cambridge.
- ALEXIADOU, A., and E. ANAGNOSTOPOULOU. (1998). 'Parameterizing Agr: Word Order, Verb-movement and EPP-checking'. *Natural Language and Linguistic Theory* 16: 491–539.
- and G. FANSELOW. (2002). 'On the Correlation between Morphology and Syntax: the Case of V-to-I', in J.W. Zwart and W. Abraham (eds.), *Studies in Comparative Germanic Syntax: Proceedings from the 15th Workshop on Comparative Germanic Syntax*. Amsterdam: John Benjamins, pp. 219–42.
- ALLEN, C. (1986). 'Reconsidering the History of *Like*'. *Journal of Linguistics* 22: 375–409.
- (1995). *Case Marking and Reanalysis: Grammatical Relations from Old to Early Modern English*. Oxford: Oxford University Press.
- ANDERSEN, H. (1973). 'Abductive and Deductive Change'. *Language* 49: 765–93.
- ANDERSON, J. (1986). 'A Note on Old English Impersonals'. *Journal of Linguistics* 22: 167–77.

- ANDERSON, S. (2002). 'Syntax and Morphology are Different: Commentary on Jonas', in D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, pp. 271–5.
- and D. LIGHTFOOT. (2002). *The Language Organ: Linguistics as Cognitive Psychology*. Cambridge: Cambridge University Press.
- ARONOFF, M. (1976). *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- ASHER, R., and T. KUMARI. (1997). *Malayalam*. London: Routledge.
- AUDI, R. (ed.). (1999). *The Cambridge Dictionary of Philosophy*. Cambridge: Cambridge University Press.
- AVRUTIN, S. (1998). 'EVENTS as Units of Discourse Representation in Root Infinitives', in J. Schaeffer (ed.), *The Interpretation of Root Infinitives and Bare Nouns in Child Language*. Occasional Papers 12. Cambridge, Mass.: MIT Press, pp. 65–91.
- AYRES-BENNETT, W. (1996). *A History of the French Language Through Texts*. London: Routledge.
- (2004). *Sociolinguistic Variation in Seventeenth-Century France: Methodology and Case Studies*. Cambridge: Cambridge University Press.
- BACH, E. (1971). 'Questions'. *Linguistic Inquiry* 2: 153–66.
- BAKER, M. (2001). *The Atoms of Language: The Mind's Hidden Rules of Grammar*. Oxford: Oxford University Press.
- BAKER, P., and C. CORNE. (1982). *Isle de France Creole: Affinities and Origins*. Ann Arbor, Mich.: Karoma.
- BAKER, M., K. JOHNSON, and I. ROBERTS. (1989). 'Passive Arguments Raised'. *Linguistic Inquiry* 20: 219–51.
- BARBER, C. (1976). *Early Modern English*. London: Andre Deutsch.
- BARBOSA, P. (1995). *Null Subjects*. PhD Dissertation, MIT.
- , E. DUARTE, and M. KATO. (2005). 'Null Subjects in European and Brazilian Portuguese'. Ms, Universities of Minho, Rio de Janeiro and Campinas.
- BARTSCH, R., and T. VENNEMANN. (1972). *Semantic Structures: A study in the Relation between Semantics and Syntax*. Frankfurt am Main: Athenäum.
- BATLLORI, M., M.-L. HERNANZ, C. PICALLO, and F. ROCA. (2005). *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press.
- BATTISTELLA, E. (1996). *The logic of markedness*. Oxford: Oxford University Press.
- and A. LOBECK. (1991). 'On Verb Fronting, Inflection Movement, and Aux Support'. *Canadian Journal of Linguistics* 36: 225–67.
- BECK, M.-L. (1998). 'L2 Acquisition and Obligatory Head-movement: English-speaking Learners of German and the Local Impairment Hypothesis'. *Studies in Second Language Acquisition* 20: 311–48.
- BEJAR, S. (2002). 'Movement, Morphology and Learnability', in D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, pp. 307–25.

- BELLETTI, A. (1990). *Generalized Verb Movement*. Turin: Rosenberg and Sellier.
- and L. RIZZI. (1988). 'Psych Verbs and θ -Theory'. *Natural Language and Linguistic Theory* 6: 291–352.
- BENNETT, P. (1979). 'Observations on the Transparency Principle'. *Linguistics* 17: 843–61.
- BENTLEY, D., and T. EYTHÓRSEN. (2003). 'Auxiliary Selection and the Semantics of Unaccusativity'. *Lingua* 114: 447–71.
- BENVENISTE, E. (1966). *Problèmes de linguistique générale*. Paris: Gallimard.
- BERGIN, O. (1934–8). 'On the Syntax of the Verb in Old Irish'. *Ériu* 12: 197–214.
- BERNSTEIN, J. (1991). 'DP's in French and Walloon: Evidence for Parametric Variation in Nominal Head Movement'. *Probus* 3: 101–26.
- 'The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain', in M. Baltin and C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 536–61.
- BERTOLO, S. (2001). *Parametric Linguistics and Learnability*. Cambridge: Cambridge University Press.
- BERWICK, R. (1985). *The Acquisition of Syntactic Knowledge*. Cambridge, Mass.: MIT Press.
- DEN BESTEN, H. (1983). 'On the Interaction of Root Transformations and Lexical Deletive Rules', in W. Abraham (ed.), *On the Formal Syntax of the Westgermania*. Amsterdam: John Benjamins, pp. 47–131.
- (1986). 'Decidability in the Syntax of Verbs of (Not Necessarily) West Germanic Languages', in *Groninger Arbeiten zur Germanistischen Linguistik* 28: 232–56. Reprinted in 1989 in Hans den Besten, *Studies in West Germanic Syntax*. Amsterdam: Rodopi, pp. 111–35.
- BHATT, R., and R. PANCHEVA. (2004). 'Late Merger of Degree Clauses'. *Linguistic Inquiry* 35: 1–45.
- BIANCHI, V. (1999). *Consequences of Antisymmetry: Headed Relative Clauses*. Berlin: Mouton de Gruyter.
- BIBERAUER, T. (2003). *Verb Second (V2) in Afrikaans: a Minimalist Investigation of Word Order Variation*. Ph.D. Dissertation, University of Cambridge.
- and M. RICHARDS. (2006). 'True Optionality: when the grammar doesn't mind', in C. Boeckx (ed.), *Minimalist Theorizing*. Amsterdam: John Benjamins, pp. 35–67.
- and I. ROBERTS. (2005a). 'Changing EPP-parameters in the history of English: accounting for variation and change', *English Language and Linguistics* 9: 5–46.
- and — (2005b). 'Cascading Parameter Changes: Internally-driven Change in Middle and Early Modern English', to appear in T. Eythórsen (ed.), *Grammatical Change and Linguistic Theory: The Rosendal Papers*.
- BICKERTON, D. (1981). *Roots of Language*. Ann Arbor, Mich.: Karoma.
- (1984). 'The Language Bioprogram Hypothesis'. *Behavioral and Brain Sciences* 7: 212–18.

- BICKERTON, D. (1988). 'Creole Languages and the Bioprogram', in F. Newmeyer (ed.), *Linguistics: The Cambridge Survey*. Cambridge: Cambridge University Press, pp. 267–84.
- (1991). 'Haunted by the Specter of Creole Genesis'. *Behavioral and Brain Sciences* 14: 354–66.
- (1999). 'How to Acquire Language without Positive Evidence: What Acquisitionists Can Learn from Creoles', in M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass.: MIT Press, pp. 49–74.
- (2004). 'Reconsidering Creole Exceptionalism'. *Language* 80: 828–33.
- BIRDSONG, D. (1992). 'Ultimate Attainment in Second Language Acquisition'. *Language* 68: 706–55.
- BLOOM, P. (1990). 'Subjectless Sentences in Child Language'. *Linguistic Inquiry* 21: 491–504.
- BOBALJK, J. (2002). 'Realizing Germanic Inflection: Why Morphology Does Not Drive Syntax'. *Journal of Comparative Germanic Linguistics* 6: 129–67.
- and D. JONAS. (1996). 'Subject Positions and the Roles of TP'. *Linguistic Inquiry* 27: 195–236.
- and H. THRÁINSSON. (1998). 'Two Heads Aren't Always Better Than One'. *Syntax* 1: 37–71.
- BOLINGER, D. (1972). *Degree Words*. The Hague: Mouton.
- BOLKESTEIN, A. (1979). 'Subject-to-Object Raising in Latin?'. *Lingua* 15–34.
- BOPP, F. (1816). *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache*. Frankfurt: Andreae, K. J. Windischmann. Reprinted Hildesheim, Olms, 1975.
- BORER, H. (1984). *Parametric Syntax*. Dordrecht: Foris.
- (1986). 'I-Subjects'. *Linguistic Inquiry* 17: 375–416.
- BORETZKY, N. (1983). *Kreolsprachen, Substrate und Sprachwandel*. Wiesbaden: Harassowitz.
- BORSLEY, R., and R. MORRIS-JONES. (2005). *Welsh Negation and Grammatical Theory*. Cardiff: University of Wales Press.
- BOŠKOVIĆ, Z. (2002). 'On Multiple *Wh*-Fronting'. *Linguistic Inquiry* 33: 351–85.
- BOWERS, J. (1993). 'The Syntax of Predication'. *Linguistic Inquiry* 24: 591–656.
- (2001) 'Predication', in M. Baltin and C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 299–333.
- BRANDI, L., and P. CORDIN. (1989). 'Two Italian Dialects and the Null Subject Parameter', in O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter*. Dordrecht: Kluwer, pp. 111–42.
- BROWN, R. (1970). *Psycholinguistics*. New York: The Free Press.
- BRUGMANN, K. (1925). *Die Syntax des einfachen Satzes im Indogermanischen*. Berlin: Walter de Gruyter.

- and B. DELBRÜCK. (1897–1916). *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen. Kurzgefaßte Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen und Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen*. Tübingen, Strasbourg.
- BRUNOT, F., and C. BRUNEAU. (1933). *Précis de grammaire historique de la langue française*. Paris: Masson.
- BURNLEY, D. (1992). *A History of the English Language through Texts*. London: Longman.
- BURZIO, L. (1986). *Italian Syntax: A Government-Binding Approach*. Dordrecht: Kluwer.
- BYBEE, J., R. PERKINS, and W. PAGLIUCA. (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- BYRNE, F. (1987). *Grammatical Relations in a Radical Creole: Verb Complementation in Saramaccan*. Amsterdam: John Benjamin.
- CALABRESE, A. (1993). 'The Sentential Complementation of Salentino: A Study of a Language without Infinitival Clauses', in A. Belletti (ed.), *Syntactic Theory and the Dialects of Italy*. Turin: Rosenberg and Sellier, pp. 28–98.
- CAMPBELL, L. (1998). *Historical Linguistics*. Edinburgh: University of Edinburgh Press.
- CANALE, M. (1978). *Word Order Change in Old English: Base Reanalysis in Generative Grammar*. PhD Dissertation, McGill University.
- CARDINALETTI, A. (1990). *Impersonal Constructions and Sentential Arguments in German*. Padua: Unipress.
- CARDINALETTI, A. (2003). Talk given at the Fourth Null-Subject Workshop, University of Cambridge.
- and L. REPETTI. (2003). 'Clitics in Northern Italian Dialects: Phonology, Syntax and Microvariation'. Talk given at the Third Workshop on Null Subjects, University of Cambridge.
- and M. STARKE. (1999). 'The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns', in H. van Riemsdijk (ed.), *Clitics in the Languages of Europe*. Berlin: de Gruyter, pp. 145–235.
- CARNIE, A. (2000). *Syntax*. Oxford: Blackwell.
- and E. GUILFOYLE. (2000). *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press.
- CAZDEN, C. (1968). 'The Acquisition of Noun and Verb Inflections'. *Child Development* 39: 433–48.
- CECCHETTO, C., and O. RENATO. (2001). 'Consequences of the Analysis of Latin Infinitival Clauses for the Theory of Case and Control'. Ms., University of Milano-Bicocca and University of Udine.
- CHENG, L. (1991). *On the Typology of Wh-Questions*. PhD Dissertation, MIT.

- CHOMSKY, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- (1959). Review of B.F. Skinner *Verbal Behavior*. *Language* 35: 26–58.
- (1964). *Current Issues in Linguistic Theory*. The Hague: Mouton.
- (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- (1966). *Cartesian Linguistics*. New York: Harper and Row.
- (1973). ‘Conditions on Transformations’, in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt, Reinhart and Winston, pp. 232–86.
- (1981) *Lectures on Government and Binding*. Dordrecht: Foris.
- (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass.: MIT Press.
- (1986). *Knowledge of Language*. New York: Praeger.
- (1993). ‘A Minimalist Program for Linguistic Theory’, in K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press, pp. 1–52.
- (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- (2000). ‘Minimalist Inquiries: The Framework’, in R. Martin, D. Michaels, and J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press, pp. 89–156.
- (2001) ‘Derivation by Phase’, in M. Kenstowicz (ed.), *Ken Hale: A Life in Language*. Cambridge, Mass.: MIT Press, pp. 1–52.
- (2002). *On Nature and Language*. Cambridge: Cambridge University Press.
- (2004). ‘Beyond Explanatory Adequacy’, in A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*. Oxford: Oxford University Press, pp. 104–31.
- (2005a). ‘Three Factors in Language Design’. *Linguistic Inquiry* 36: 1–22.
- (2005b). ‘Biolinguistics and the Human Capacity’, lecture given at MTA, Budapest, May 17, 2004, in N. Chomsky (2006). *Language and mind*, 3rd edn. Cambridge: Cambridge University Press, pp. 173–85.
- (2005c). ‘On Phases’. Ms., MIT.
- and M. HALLE. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- CHUNG, S. (2005). ‘What Fronts? On the VP Raising Account of Verb-Initial Order’, in A. Carnie, H. Harley, and S. Dooley (eds.), *Verb-First*, Amsterdam: John Benjamins, pp. 9–29.
- CINQUE, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- (2004). ‘“Restructuring” and Functional Structure’, in A. Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3*. Oxford: Oxford University Press, pp. 132–91.
- CLAHSEN, H., and U. HONG. (1995). ‘Agreement and Null Subjects in German L2 Development: New Evidence from Reaction-time Experiments’. *Second Language Research* 11: 57–87.

- and M. PENKE. (1992). 'The Acquisition of Agreement Morphology and its Syntactic Consequences: New Evidence on German Child Language from the Simone-corpus', in J. Meisel (ed.), *The Acquisition of Verb Placement*. Dordrecht: Kluwer, pp. 181–224.
- and K. SMOLKA. (1995). 'Psycholinguistic Evidence and the Description of V2 Phenomena in German', in H. Haider and M. Prinzhorn (eds.), *Verb Second Phenomena in the Germanic Languages*. Dordrecht: Foris.
- , C. KURSAWE, and M. PENKE. (1995). 'Introducing CP: *Wh*-questions and subordinate clauses in German child language', in *Essex Research Reports in Linguistics*, 7: 1–28. University of Essex, Department of Linguistics.
- CLARK, R., and I. ROBERTS. (1993). 'A Computational Approach to Language Learnability and Language Change'. *Linguistic Inquiry* 24: 299–345.
- COCCHI, G. (1995). *La selezione dell'ausiliare*. Padua: Unipress.
- COLLINS, C. (2005). 'A Smuggling Approach to the Passive in English'. *Syntax* 8: 81–120.
- COMRIE, B. (1989). *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- , S. MATTHEWS, and M. POLINSKY. (2003). *The Atlas of Languages*. New York: Quarto.
- COOK, V., and M. NEWSON. (1996). *Chomsky's Universal Grammar: An Introduction*. Oxford: Blackwell.
- COPPIETERS, R. (1987). 'Comparative Differences between Native and Near-native Speakers'. *Language* 63: 544–73.
- COTTELL, S. (2002). *The Comparative Syntax of Cleft Constructions*. PhD Dissertation, University of Wales, Bangor.
- CRAIN, S., and P. PIETROSKI. (2002). Reply to Pullum and Scholz. *The Linguistic Review* 19: 163–84.
- CROFT, W. (2000). *Explaining Language Change*. London: Longman.
- (2003). *Typology and Universals*. Cambridge: Cambridge University Press.
- DEGRAFF, M. (1992). 'The Syntax of Predication in Haitian', in K. Broderick (ed.), *Proceedings of NELS 22*. GLSA, Amherst, Mass., pp. 103–17.
- (1993). 'A Riddle on Negation in Haitian'. *Probus* 5: 63–93.
- (1994). 'To Move or Not to Move? Placement of Verbs and Object Pronouns in Haitian and in French', in K. Beals *et al.* (eds.), *Papers from the 30th Meeting of the Chicago Linguistics Society*, Chicago Linguistics Society, Chicago, pp. 141–155.
- (1999). *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass.: MIT Press.
- (2003). 'Against Creole Exceptionalism'. *Language* 79: 391–410.
- (2004). 'Against Creole Exceptionalism (redux)'. *Language* 80: 834–9.
- (2005). 'Morphology and Word Order in "Creolization" and Beyond', in G. Cinque and R. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press, pp. 293–372.

- DEGRAFF, M. (to appear). 'Language Acquisition in Creolization: Uniformitarian Approaches', in S. Kouwenberg and J. Singler (eds.), *Handbook of Pidgin and Creole Linguistics*. Oxford: Blackwell.
- and Y. DEJEAN. (1994). 'On Haitian Creole's "Very Strict" Adjacency Principle'. Paper given at the meeting of the Society for Pidgin and Creole Linguistics, Boston.
- DELBRÜCK, B. (1893–1900). *Vergleichenden Syntax der indogermanischen Sprachen*. In K. Brugmann and B. Delbrück (1897–1916), *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen. Kurzgefaßte Darstellung der Geschichte des Altindischen, Altiranischen (Avestischen und Altpersischen), Altarmenischen, Altgriechischen, Albanesischen, Lateinischen, Oskisch-Umbrischen, Altirischen, Gotischen, Althochdeutschen, Litauischen und Altkirchenslavischen*. Tübingen, Strasbourg, V, Vols 1–3.
- DELL, F. (1985). *Les règles et les sons: introduction à la phonologie generative*. Paris: Hermann.
- DELSING, L.-O. (2000). 'From OV to VO in Swedish,' in S. Pintzuk, G. Tsoulas, and A. Warner (eds.), *Diachronic Syntax: Models and Mechanisms*. Oxford: Oxford University Press, pp. 255–74.
- DENISON, D. (1985). 'The Origins of Periphrastic *Do*: Ellegård and Visser Reconsidered', in R. Eaton *et al.* (eds.), *Papers from the 4th International Conference on Historical Linguistics, Amsterdam, April 10–13, 1985*. Amsterdam: John Benjamins, pp. 45–60.
- (1990). Auxiliary + Impersonal in Old English. *Folia Linguistica Historica* 9: 139–66.
- (1993) *English Historical Syntax*. London: Longman.
- (1999) 'Slow, Slow, Quick, Quick, Slow: the Dance of Language Change?', in A. Bringas López *et al.* (eds.), '*Woonderous Ænglissce*': *SELIM Studies in Medieval English Language*, Vigo: Universidade de Vigo (Servicio de Publicacións), pp. 51–64.
- (2003). 'Log(ist)ic and Simplistic S-curves', in R. Hickey (ed.), *Motives for language change*. Cambridge: Cambridge University Press, pp. 54–70.
- DÉPREZ, V. (1997). 'Two Types of Negative Concord'. *Probus* 9: 103–43.
- (1999). 'The Roots of Negative Concord in French and French-Lexicon Creoles', in M. DeGraff (ed.), (1999) *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass.: MIT Press, pp. 375–428.
- (2000). 'Parallel (A)symmetries and the Internal Structure of Negative Expressions'. *Natural Language and Linguistic Theory* 18: 253–342.
- DERRICK MCCLURE, J. (1994). 'English in Scotland', in R. Burchfield (ed.), *The Cambridge History of the English Language Volume V: English in Britain and Overseas, Origins and Development*. Cambridge: Cambridge University Press, pp. 23–93.

- DOHERTY, C. (2000a). 'Verb Second and Tmesis in Early Irish', in S. Chang *et al.* (eds.), *Proceedings of the Twenty-Fifth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, Ca.: Berkeley Linguistics Society, pp. 98–108.
- (2000b). 'Residual Verb Second in Early Irish: On the Nature of Bergin's Construction'. *Diachronica* 17: 5–38.
- DRESHER, E. (1999). 'Charting the Learning Path: Cues to Parameter Setting'. *Linguistic Inquiry* 30: 27–68.
- and J. KAYE. (1990). 'A Computational Learning Model for Metrical Phonology'. *Cognition* 34: 137–95.
- DRYER, M. (1992). 'On the Greenbergian Word-Order Correlations'. *Language* 68: 81–138.
- (2005a). 'Order of Object and Verb', in M. Haspelmath, M. Dryer, D. Gil, and B. Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, pp. 338–41.
- (2005b). 'Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase', in M. Haspelmath, M. Dryer, D. Gil, and B. Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, pp. 386–9.
- DUARTE, E. (1993). 'Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil', in I. Roberts and M. Kato (eds.), *Viagem Diacrônica pelas Fases do Português Brasileiro: Homagem a Fernando Tarallo*, Campinas: Editora da Unicamp, pp. 107–28.
- (1995). *A Perda do Princípio 'Evite pronome' no Português Brasileiro*. PhD Dissertation, Unicamp.
- DUPUIS, F. (1988). 'Pro-drop dans les subordinnées en ancien français', in P. Hirschbuhler and A. Rochette (eds.), *Aspects de la syntaxe historique du français*, *Revue québécoise de linguistique théorique et appliquée* 7: 41–62.
- (1989). *L'expression du sujet dans les subordinnées en ancien français*. PhD Dissertation, Université de Montréal.
- DURIE, M., and M. ROSS. (1996). *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. Oxford: Oxford University Press.
- DURRLEMAN, S. (2004). 'Topics, Focus and wh in Jamaican Creole'. Handout of talk given at the Society for Pidgin and Creole Linguistics Conference, January 9, 2004.
- EINHORN, E. (1974). *Old French*. Cambridge: Cambridge University Press.
- ELDRIDGE, N., and S. GOULD. (1972). 'Punctuated Equilibria: An Alternative to Phyletic Gradualism', in T. Scopf (ed.), *Models in Paleobiology*. San Francisco: Freeman, pp. 82–115.
- ELLEGÅRD, A. (1953). *The Auxiliary Do: The Establishment and Regulation of its Use in English*. Stockholm: Almqvist and Wiksell.
- EMONDS, J. (1978). 'The Verbal Complex V-V' in French'. *Linguistic Inquiry* 9: 151–75.

- EMONDS, J. (1980). 'Word Order in Generative Grammar'. *Journal of Linguistic Research* 1: 33–54.
- ERNOUT, A., and F. THOMAS. (1993). *Syntaxe Latine*. Paris: Klincksieck.
- ERNST, G. (1985). *Gesprochenes Französisch zu Beginn des 17. Jahrhunderts. Direkte Rede in Jean Héroards 'Histoire particulière de Louis XIII' (1605–1610)*. Tübingen: Niemeyer.
- EYTHÖRSSEN, T., and J. BARÖDAL. (2005). 'Oblique Subjects: A Common Germanic Inheritance'. *Language* 81: 824–81.
- FAARLUND, J.-T. (1994). 'Old and Middle Scandinavian', in E. König and J. van der Auwera (eds.), *The Germanic Languages*. London: Routledge, pp. 38–71.
- (2004a). 'Ancient Nordic', in R. Woodard (ed.), *The Cambridge Encyclopedia of Ancient Languages*. Cambridge: Cambridge University Press, pp. 907–21.
- (2004b). *Syntax of Old Norse: With a Survey of the Inflectional Morphology and a Complete Bibliography, A Grammar of Old Norse*. Oxford: Blackwell.
- FARKAS, D. (1992). 'On the Semantics of Subjunctive Complements', in P. Hirschbühler and K. Koerner (eds.), *Romance Linguistics and Modern Linguistic Theory*. Amsterdam: John Benjamins, pp. 67–104.
- FELSENSTEIN, J. (2001). *PHYLIP: Phylogeny Inference Package. Version 3.6*, Department of Genetics, University of Washington.
- FERGUSON, C. (1959). 'Diglossia'. *Word* 15: 325–40.
- FISCHER, O. and F. VAN DER LEEK (1983) 'The Demise of the Old English Impersonal Construction'. *Journal of Linguistics* 19: 337–68.
- , B. LOS, A. VAN KEMENADE, and W. VAN DER WURFF. (2000). *The Syntax of Early English*. Cambridge: Cambridge University Press.
- FLYNN, S. (1996). 'A Parameter-setting Approach to Second Language Acquisition', in W. Ritchie and T. Bhatia (eds.), *Handbook of Language Acquisition*. San Diego: Academic Press, pp. 121–58.
- and G. MARTOHARDJONO (1994). 'Mapping from the Initial State to the Final State: the Separation of Universal Principles and Language-specific Principles', in B. Lust, M. Suñer, and J. Whitman (eds.), *Syntactic Theory and First-language Acquisition: Cross-linguistic Perspectives, Volume 1: Heads, Projections and Learnability*. Hillsdale, NJ: Lawrence Erlbaum, pp. 319–35.
- FODOR, J. (1998). 'Unambiguous triggers'. *Linguistic Inquiry* 29: 1–36.
- and C. CROWTHER (2002). 'Understanding Stimulus Poverty Arguments'. *The Linguistic Review* 19: 105–46.
- FONTAINE, C. (1985). *Application de méthodes quantitatives en diachronie: l'inversion du sujet en français*. MA Thesis, Université de Québec à Montréal.
- FONTANA, J.-M. (1993). *Phrase Structure and the Syntax of Clitics in the History of Spanish*. PhD Dissertation, University of Pennsylvania.
- FORTSON, B. (2004). *Indo-European Language and Culture*. Oxford: Blackwell.
- FOSTER, T. and W. VAN DER WURFF. (1997). 'From Syntax to Discourse: the Function of Object-Verb Word Order in Late Middle English', in J. Fisiak (ed.), *Studies in Middle English Linguistics*. Berlin: Mouton de Gruyter, pp. 135–56.

- FOULET, L. (1921). 'Comment ont évolué les formes de l'interrogation?' *Romania* 47: 243–348.
- (1990). *Petite syntaxe de l'ancien français*. Paris: Champion.
- FOWLER, J. (1986). 'The Social Stratification of (r) in New York City Department Stores, 24 Years after Labov'. Ms., New York University.
- FOX, A. (1995). *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press.
- FRANZÉN, T. (1939). *La syntaxe des pronoms personnels sujet en ancien français*. Uppsala: Almqvist.
- FREEZE, R. (1992). 'Existentials and Other Locatives'. *Language* 68: 553–95.
- FRIEDEMANN, M.-A., and L. RIZZI. (2000). *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics*. London: Longman.
- FRIEDRICH, P. (1975). *Proto-Indo-European Syntax; The Order of Meaningful Elements*. Butte: Montana College of Mineral Science and Technology.
- FUß, E., and C. TRIPS. (2002). 'Variation and Change in Old and Middle English. On the Validity of the Double Base Hypothesis'. *Journal of Comparative Germanic Linguistics* 4: 171–224.
- VAN DER GAAF, W. (1904). *The Transition from the Impersonal to the Personal Construction in Middle English*. *Anglistische Forschungen* 14. Reprinted 1967. Amsterdam: Swets and Zeitlinger.
- GAMILLSCHEG, E. (1957). *Historische französische Syntax*. Tübingen: Niemeyer.
- GARRETT, A. (1990). *The Syntax of Anatolian Pronominal Clitics*. PhD Dissertation, Harvard University.
- GAVARRÒ, A., and J. SOLÀ. (2004). 'Wh-subextraction in Child Catalan'. Paper presented at the 2nd Lisbon Meeting on Language Acquisition.
- GEORGOPOULOS, C. (1991). *Syntactic Variables: Resumptive Pronouns and A' Binding in Palauan*. Dordrecht: Kluwer.
- GERKEN, L. (1991). 'The Metrical Basis of Children's Subjectless Sentences'. *Journal of Memory and Language* 30: 431–51.
- GIANNAKIDOU, A. (1997). *The Landscape of Polarity Items*. Groningen Dissertations in Linguistics, 18.
- (1998). *Polarity Sensitivity as (Non)veridical Dependency*. Amsterdam: John Benjamins.
- (2000). Negative ... Concord? *Natural Language and Linguistic Theory* 18: 457–523.
- GIANOLLO, C., C. GUARDIANO, and G. LONGOBARDI. (2004). 'Historical Implications of a Formal Theory of Syntactic Variation'. Paper presented at the 8th Diachronic Generative Syntax Conference, Yale University.
- GIL, D. (2005). 'Numeral Classifiers', in M. Haspelmath, M. Dryer, D. Gil, and B. Comrie (eds.), *The World Atlas of Language Structures*. Oxford: Oxford University Press, pp. 226–9.

- GIORGI, A., and F. PIANESI. (1997). *Tense and Aspect: From Semantics to Morphosyntax*. Oxford: Oxford University Press.
- GOLDIN-MEADOW, S. (2005). 'What Language Creation in the Manual Modality Tells Us about the Foundations of Language'. *The Linguistic Review* 22: 199–226.
- GÖRLACH, M. (1991). *Introduction to Early Modern English*. Cambridge: Cambridge University Press.
- GRAY, D. (1985). *The Oxford Book of Late Medieval Prose and Verse*. Oxford: Oxford University Press.
- GREEN, J. (1988). 'Romance Creoles', in M. Harris and N. Vincent (eds.), *The Romance Languages*. London: Routledge, pp. 420–74.
- GREENBERG, J. (1963). 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements', in J. Greenberg (ed.), *Universals of Grammar*. Cambridge, Mass.: MIT Press, pp. 73–113.
- (1980). 'Circumfixes and typological change', in E. Traugott, R. Labrum, and S. Shepherd (eds.), *Papers from the 14th International Conference on Historical Linguistics*, Amsterdam: John Benjamins, pp. 233–41.
- GRÉGOIRE, A. (1937–47). *L'apprentissage du langage*. Liège: Droz.
- GUARDIANO, C., and G. LONGOBARDI. (2003). 'Parametric Syntax as a Source of Historical-comparative Generalisations'. Ms., University of Trieste.
- and — (2005). 'Parametric Comparison and Language Taxonomy', in M. Batllori, M.-Ll. Hernanz, C. Picallo, and F. Roca (eds.), *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press, pp. 149–74.
- GUASTI, M.-T. (1991). *Causative and Perception Verbs: A Comparative Study*. Turin: Rosenberg and Sellier.
- (1996). 'The Acquisition of Italian Interrogatives', in H. Clahsen (ed.), *Generative Perspectives on Language Acquisition*. Amsterdam: John Benjamins, pp. 241–69.
- (2002). *Language Acquisition: The Growth of Grammar*. Cambridge, Mass.: MIT Press.
- HAEBERLI, E. (2002). *Features, Categories and the Syntax of A-Positions. Cross-Linguistic Variation in the Germanic Languages*. Kluwer: Dordrecht.
- HAEGEMAN, L. (1994). *Introduction to Government and Binding Theory*. Oxford: Blackwell.
- (1995a). 'Root Infinitives and Initial Root Null Subjects in Early Dutch', in C. Koster and F. Wijnen (eds.), *Proceedings of the Groningen Assembly on Language Acquisition*. Groningen: University of Groningen Press.
- (1995b). 'Root Infinitives, Tense and Truncated Structures in Dutch'. *Language Acquisition* 4: 205–55.
- (2000). 'Adult Null Subjects in Non Pro-drop Languages', in M.-A. Friedemann and L. Rizzi (2000), *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics*. London: Longman, pp. 129–69.

- (2005). *Thinking Syntactically: A Guide to Argumentation and Analysis*. Oxford: Blackwell.
- HALE, M. (1995). 'Wackernagel's Law in the Rigveda'. Ms., University of Concordia.
- (1996). 'Theory and Method in Historical Linguistics'. Ms., University of Concordia.
- (1998). 'Diachronic Syntax'. *Syntax* 1: 1–18.
- HALLE, M. (1962). 'Phonology in a Generative Grammar'. *Word* 18: 54–72.
- HAMANN, C., and K. PLUNKETT. (1998). 'Subjectless Sentences in Child Danish'. *Cognition* 69: 35–72.
- L. RIZZI, and U. FRAUENFELDER. (1996). 'On the Acquisition of the Pronominal System in French'. *Recherches Linguistiques* 24: 83–101.
- HANCOCK, I. (1979). 'The Relationship of Black Vernacular English to the Atlantic Creoles'. Working paper of the African and Afro-American Studies and Research Center, University of Texas at Austin.
- (1985). 'A Preliminary Sketch of Trinidad Creole French'. *Amsterdam Creole Studies* 8: 27–40.
- HARRIS, A., and L. CAMPBELL. (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- HARRIS, M. (1978). *The Evolution of French Syntax: A Comparative Approach*. London: Longman.
- HASPELMATH, M. (1989). 'From Purposive to Infinitive – a Universal Path of Grammaticalization'. *Folia Linguistica Historica* 10: 287–310.
- (1997). *Indefinite Pronouns*. Oxford: Oxford University Press.
- (1998). 'Does Grammaticalization Need Reanalysis?'. *Studies in Language* 22: 315–51.
- M. DRYER, D. GIL, and B. COMRIE (eds.), (2005) *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- HAUSER, M., N. CHOMSKY, and W. FITCH. (2002). 'The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?'. *Science* 298: 1569–79.
- HAWKINS, J. (1983). *Word Order Universals*. New York: Academic Press.
- HAWKINS, R. (2001). *Second Language Syntax: A Generative Introduction*. Oxford: Blackwell.
- HEINE, B., U. CLAUDI, and F. HÜNNEMEYER. (1991). *Grammaticalization: a Conceptual Framework*. Chicago: University of Chicago Press.
- and T. KUTEVA. (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- and M. REH. (1984) *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.
- , T. GÜLDEMANN, C. KILIAN-HATZ, D. LESSAU, H. ROBERG, M. SCHLADT, and T. STOLZ. (1993). *Conceptual Shift: a Lexicon of Grammaticalization Processes in African Languages*. University of Cologne: Institut für Afrikanistik (Afrikanistische Arbeitspapiere 34/35).

- HENRY, A. (1995). *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.
- HILTUNEN, R. (1983). *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb*. Turku: Turun Yliopisto.
- HINTERHÖLZL, R. (1997). 'An XP-movement account of restructuring'. Ms., University of Southern California.
- HIRSCHBUHLER, P. (1990). 'La légitimation de la construction V1 à sujet nul dans la prose et le vers en ancien français'. *Revue québécoise de linguistique* 19: 32–55.
- and M.-O. JUNKER (1988). 'Remarques sur les sujets nuls en subordonnées en ancien et en moyen français', in P. Hirschbuhler and A. Rochette (eds.), *Aspects de la syntaxe historique du français, Revue québécoise de linguistique théorique et appliquée* 7: 63–84.
- HOCK, H. (1991). *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
- and B. JOSEPH. (1996). *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin: Mouton de Gruyter.
- HOCKETT, C. (1958). *A Course in Modern Linguistics*. New York: Macmillan.
- HOEKSEMA, J. (2000). 'Negative Polarity Items: Triggering, Scope and C-Command', in L. Horn and Y. Kato (eds.), *Negation and Polarity. Semantic and Syntactic Perspectives*. Oxford: Oxford University Press, pp. 115–46.
- HOEKSTRA, T., and N. HYAMS. (1998). 'Aspects of Root Infinitives'. *Lingua* 106: 81–112.
- HOGG, R. (1992). 'Phonology and Morphology', in R. Hogg (ed.), *The Cambridge History of the English Language Volume I: The Beginnings to 1066*. Cambridge: Cambridge University Press, pp. 67–167.
- (1992–2001). *The Cambridge History of the English Language Volumes I–VI*. Cambridge: Cambridge University Press.
- HOLM, J. (1988). *Pidgins and Creoles, Volumes I–II*. Cambridge: Cambridge University Press.
- HOLMBERG, A. (1986). *Word Order and Syntactic Features in Scandinavian Languages and English*. PhD Dissertation, University of Stockholm.
- (1999). 'Remarks on Holmberg's Generalization'. *Studia Linguistica* 53: 1–39.
- (2005). 'Is There a Little Pro? Evidence from Finnish', *Linguistic Inquiry* 36: 533–64.
- , C. JOHNS, A. NAYUDU, and M. SHEEHAN. (2005). 'Partial Null-Subject Languages.' Ms., University of Newcastle-Upon-Tyne.
- HOPPER, P., and E. TRAUGOTT. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.
- HORN, L. (2000). 'Pick a Theory (Not Just Any Theory): Indiscriminatives and the Free-Choice Indefinite', in L. Horn and Y. Kato (eds.), *Negation and Polarity: Semantic and Syntactic Perspectives*. Oxford: Oxford University Press, pp. 147–92.

- and Y. KATO. (2000). *Negation and Polarity: Semantic and Syntactic Perspectives*. Oxford: Oxford University Press.
- HORNSTEIN, N., and A. WEINBERG. (1981). 'Case Theory and Preposition Stranding'. *Linguistic Inquiry* 12: 55–92.
- , J. NUNES, and K. GROHMANN. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.
- HORROCKS, G. (1997). *Greek: A History of the Language and Its Speakers*. London: Longman.
- HRÓARSDÓTTIR, T. (1996). 'The Decline of OV Word Order in the Icelandic VP: A Diachronic Study'. *Working Papers in Scandinavian Syntax* 57: 92–141, Department of Scandinavian Languages, University of Lund.
- (1999). *Verb Phrase Syntax in the History of Icelandic*. PhD Dissertation, University of Tromsø.
- (2000). *Word Order Change in Icelandic: From OV to VO*. Amsterdam: John Benjamins.
- HUANG, J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. PhD Dissertation, MIT.
- (1984). 'On the Distribution and Reference of Empty Pronouns'. *Linguistic Inquiry* 15: 531–74.
- (1989). 'Pro-drop in Chinese: A Generalized Control Theory', in O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter*. Dordrecht: Kluwer, pp. 185–214.
- HYAMS, N. (1986). *Language Acquisition and the Theory of Parameters*. Dordrecht: Kluwer.
- (1992). 'A Reanalysis of Null Subjects in Child Language', in J. Weissenborn, H. Goodluck, and T. Roeper (eds.), *Theoretical Issues in Language Acquisition*. New Jersey: Lawrence Erlbaum, pp. 249–67.
- (1996). 'The Underspecification of Functional Categories in Early Grammar', in H. Clahsen (ed.), *Generative Perspectives on Language Acquisition*. Amsterdam: John Benjamins, pp. 91–128.
- and K. WEXLER. (1993). 'The Grammatical Basis of Null Subjects in Child Language'. *Linguistic Inquiry* 24: 421–60.
- INGHAM, R. (2001). 'The Structure and Function of Expletive *There* in Pre-modern English'. *Reading Working Papers in Linguistics* 5: 231–49.
- (2002). 'Negated Subject and Object Positions in 15th century Non-literary English'. *Language Variation and Change* 14: 291–322.
- JACKENDOFF, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- and S. PINKER. (2005). 'The Nature of the Language Faculty and its Implications for Evolution of Language (Reply to Fitch, Hauser, and Chomsky)'. *Cognition* 97: 211–25.
- JAEGGLI, O. (1982). *Topics in Romance Syntax*. Dordrecht: Foris.

- JAKOBSON, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze*. English translation: *Child Language, Aphasia and Phonological Universals*, A. Kieler (translator). The Hague: Mouton.
- JASANOFF, J. (2004). 'Gothic', in R. Woodard (ed.), *The Cambridge Encyclopedia of Ancient Languages*. Cambridge: Cambridge University Press, pp. 881–906.
- JESPERSEN, O. (1909–49). *A Modern English Grammar on Historical Principles I–VII*. London/Copenhagen: Allen and Unwin.
- (1917). *Negation in English and Other Languages*. Copenhagen: Det Kgl. Danske Videnskabernes Selskab. Historisk-filologiske Meddelelser 1.1–151.
- (1938). *Growth and Structure of the English Language*. Oxford: Blackwell.
- JONAS, D. (1996). *Clause Structure and Verb Syntax in Scandinavian and English*. PhD Dissertation, Harvard University.
- (2002). 'Residual V-to-I', in D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, pp. 251–70.
- JONES, C. (1997). *The Edinburgh History of the Scots Language*. Edinburgh: Edinburgh University Press.
- JONES, M. (1996). *Foundations of French Syntax*. Cambridge: Cambridge University Press.
- JOUITTEAU, M. (2005). *La syntaxe comparée du Breton*. PhD Dissertation, University of Nantes.
- KALLEN, J. (1994). 'English in Ireland', in R. Burchfield (ed.), *The Cambridge History of the English Language Volume V: English in Britain and Overseas, Origins and Development*. Cambridge: Cambridge University Press, pp. 148–96.
- KASTOVSKY, D. (1992). 'Semantics and Vocabulary', in R. Hogg (ed.), *The Cambridge History of the English Language Volume I: The Beginnings to 1066*. Cambridge: Cambridge University Press, pp. 290–408.
- KATO, M. (2000). 'The Partial Prodrop Nature and the Restricted VS Order in Brazilian Portuguese', in M. Kato and E. Negrão (eds.), *The Null Subject Parameter in Brazilian Portuguese*. Frankfurt: Vervuert-IberoAmericana, pp. 223–58.
- and E. NEGRÃO (eds.). *The Null Subject Parameter in Brazilian Portuguese*. Frankfurt: Vervuert-IberoAmericana.
- KAYNE, R. (1975). *French Syntax: The Transformational Cycle*. Cambridge, Mass.: MIT Press.
- (1983). 'Chains, Categories External to S, and French Complex Inversion'. *Natural Language and Linguistic Theory* 1: 107–139. Reprinted in Kayne (1984).
- (1984). *Connectedness and Binary Branching*. Dordrecht: Foris.
- (1989). 'Null Subjects and Clitic-Climbing', in O. Jaeggli and K. Safir (eds.), *The Null Subject Parameter*. Dordrecht: Kluwer, pp. 239–61.
- (1991). 'Romance Clitics, Verb Movement and PRO'. *Linguistic Inquiry* 22: 647–86.
- (1993). 'Towards a Modular Theory of Auxiliary Selection'. *Studia Linguistica* 47: 3–31.

- (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- (2000). *Parameters and Universals*. Oxford: Oxford University Press.
- and J.-Y. POLLOCK. (1978). 'Stylistic Inversion, Successive Cyclicity and Move NP in French'. *Linguistic Inquiry* 9: 595–621.
- — (2001). 'New Thoughts on Stylistic Inversion', in A. Hulk and J.-Y. Pollock (eds.), *Subject Inversion in Romance and the Theory of Universal Grammar*. Oxford: Oxford University Press, pp. 107–62.
- KEENAN, E. (2002). 'Explaining the Creation of Reflexive Pronouns in English,' in D. Minkova and R. Stockwell (eds.), *Studies in the History of English: A Millennial Perspective*. Berlin: Mouton de Gruyter, pp. 325–55.
- KEGL, J. (to appear). 'The Case of Signed Languages in the Context of Pidgin and Creole Studies'. in S. Kouwenberg and J. Singler (eds.), *Handbook of Pidgin and Creole Linguistics*. Oxford: Blackwell.
- A. SENGHAS, and M. COPPOLA. (1999). 'Creation through Contact: Sign Language Emergence and Sign Language Change in Nicaragua', in M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass.: MIT Press, pp. 179–238.
- VAN KEMENADE, A. (1987). *Syntactic Case and Morphological Case in the History of English*. Dordrecht: Foris.
- (2000). 'Jespersen's Cycle Revisited: Formal Properties of Grammaticalization', in S. Pintzuk, G. Tsoulas, and A. Warner (eds.), *Diachronic Syntax: Models and Mechanisms*. Oxford: Oxford University Press, pp. 51–74.
- KENNEDY, B. (1962). *The Revised Latin Primer*. London: Longman.
- KENSTOWICZ, M. (1991). *Phonology in Generative Grammar*. Oxford: Blackwell.
- KIHM, A. (1994). *Kriyol Syntax: The Portuguese-based Creole Language of Guinea-Bissau*. Amsterdam: Benjamins.
- KING, R. (2000). *The Lexical Basis of Grammatical Borrowing: A Prince Edward Island Case Study*. Amsterdam: John Benjamins.
- KIPARSKY, P. (1973). '“Elsewhere” in Phonology', in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt, Reinhart and Winston, pp. 93–106.
- (1995). 'Indo-European Origins of Germanic Syntax', in A. Battye and I. Roberts (eds.), *Clause Structure and Language Change*. Oxford: Oxford University Press, pp. 140–69.
- (1996). 'The Shift to Head-Initial VP in Germanic', in H. Thráinsson, S. Epstein, and S. Peter (eds.), *Studies in Comparative Germanic Syntax II*. Dordrecht: Kluwer, pp. 140–79.
- (2003). 'The Phonological Basis of Sound Change,' in R. Janda and B. Joseph (eds.), *Handbook of Historical Linguistics*. Oxford: Blackwell, pp. 313–342 Reprinted from J. Goldsmith (ed.). 1995 *Handbook of Phonological Theory*. Oxford: Blackwell.

- KIPARSKY, P. and C. KIPARSKY (1971). 'Fact', in L. Jakobovits and D. Steinberg (eds.), *Semantics: An Interdisciplinary Reader*. Cambridge University Press, pp. 345–69.
- DE KOK, A. (1985). *La place du pronom régime conjoint en français: une étude diachronique*. Amsterdam: Rodopi.
- KOOPMAN, H. (1984). *The Syntax of Verb-movement: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.
- (1986). 'The Genesis of Haitian: Implications of a Comparison of Some Features of the Syntax of Haitian, French and West African Languages', in P. Muysken and N. Smith (eds.), *Substrata vs. Universals in Creole Languages*. Amsterdam: John Benjamin, pp. 231–58.
- and A. SZABOLSCI. (2000). *Verbal Complexes*. Cambridge, Mass.: MIT Press.
- KORNFILT, J. (1997). *Turkish*. London: Routledge.
- KOSTER, J. (1975). 'Dutch as an SOV Language'. *Linguistic Analysis* 1: 111–36.
- (2000). 'Pied Piping and The Word Orders of English and Dutch', In: M. Hirotani, A. Coetzee, N. Hall, and J.-Y. Kim (eds.), *NELS 30: Proceedings of the North East Linguistic Society*. GLSA, Amherst, MA, pp. 415–26.
- KOUWENBERG, S. (1990). 'Complementizer *pa*, the Finiteness of its Complements and Remarks on Empty Categories in Papiamentu'. *Journal of Pidgin and Creole Languages* 5: 39–51.
- (1994). 'Berbice Dutch', in J. Arends, P. Muysken, and N. Smith (eds.), *Pidgins and Creoles*, Amsterdam: John Benjamins, pp. 233–246.
- KROCH, A. (1989). 'Reflexes of Grammar in Patterns of Language Change'. *Language Variation and Change* 1: 199–244.
- (1994). 'Morphosyntactic Variation', in Beals *et al.* (eds.), *Proceedings of the Thirtieth Annual Meeting of the Chicago Linguistic Society*, Vol. 2, pp. 180–201.
- (2000). 'Syntactic Change', in M. Baltin and C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 629–739.
- and A. TAYLOR. (1997). 'Verb Movement in Old and Middle English: Dialect Variation and Language Contact', in A. van Kemenade and N. Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, pp. 297–325.
- (2000). 'Verb-complement Order in Middle English', in S. Pintzuk, G. Tsoulas, and A. Warner (eds.), *Diachronic Syntax: Models and Mechanisms*. Oxford: Oxford University Press, pp. 132–63.
- and D. RINGE. (2000). 'The Middle English Verb-second Constraint: A Case Study in Language Contact and Language Change', in S. Herring, L. Schoesler, and P. van Reenen (eds.), *Textual Parameters in Older Language*. Amsterdam: John Benjamins (pp. 353–91).
- KÜHNER, R., and C. STEGMANN. (1955). *Ausführliche Grammatik der Lateinischen Sprache*. Leverkusen: Gottschalksche Verlagsbuchhandlung.

- LABOV, W. (1966). *The Social Stratification of Speech in New York City*. Washington D.C.: Center for Applied Linguistics.
- (1972). *Language in the Inner City: Studies in the Black English Vernacular*. New York: Academic Press.
- (1994). *Principles of Language Change: Internal Factors*. Oxford: Blackwell.
- LADUSAW, W. (1980). *Polarity Sensitivity as Inherent Scope Relations*. New York: Garland.
- LAKOFF, R. (1972). 'Another Look at Drift', in R. Stockwell and R. Macauley (eds.), *Linguistic Change and Generative Theory*. Bloomington: Indiana University Press, pp. 172–98.
- LARSON, R. (1988). 'On the Double-Object Construction'. *Linguistic Inquiry* 19: 335–91.
- LASNIK, H. (1983). *Essays on Restrictiveness and Learnability*. Dordrecht: Kluwer.
- J. URIAGEREKA, and C. BOECKX. (2004). *A Course in Minimalist Syntax*. Oxford: Blackwell.
- LASS, R. (1992). 'Phonology and Morphology', in N. Blake (ed.), *The Cambridge History of the English Language Volume II: 1066–1476*. Cambridge: Cambridge University Press, pp. 23–155.
- (1997). *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- (1999). 'Phonology and Morphology', in R. Lass (ed.), *The Cambridge History of the English Language Volume III: 1476–1776*. Cambridge: Cambridge University Press, pp. 56–186.
- LAW, V. (2003). *The History of Linguistics in Europe: From Plato to 1600*. Cambridge: Cambridge University Press.
- LÉAUTAUD (1989). *Le Fléau. Journal Particulier. 1917–1939*. Paris: Mercure de France.
- LEDGEWAY, A. (1998). 'Variation in the Romance Infinitive: The Case of the Southern Calabrian Inflected Infinitive'. *Transactions of the Philological Society* 96: 1–61.
- (2000). *A Comparative Syntax of the Dialects of Southern Italy: A Minimalist Approach*. Oxford: Blackwell.
- LEE, F. (2000). 'VP Remnant Movement and VSO in Quiavini Zapotec', in A. Carnie and E. Guilfoyle (eds.), *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press, pp. 143–62.
- LEFEBVRE, C. (1998). *Creole Genesis and the Acquisition of Grammar: The Case of Haitian Creole*. Cambridge: Cambridge University Press.
- and J. LUMSDEN. (1989). 'Les langues créoles et la théorie linguistique'. *Canadian Journal of Linguistics; Numéro Spécial: La créolisation* 34: 249–72.
- LEHMANN, C. (1986). 'Grammaticalization and Linguistic Typology'. *General Linguistics* 26: 3–22.
- (1995). *Thoughts on Grammaticalization*. Munich: Lincom Europa.
- LEHMANN, W. (1973). 'A Structural Principle of Language and its Implications'. *Language* 49: 47–66.

- LEHMANN, W. (1993). *Theoretical Bases of Indo-European Linguistics*. London: Routledge.
- LENNEBERG, E. (1967). *Biological Foundations of Language*. New York: John Wiley and Sons.
- LEOPOLD, W. (1939–49). *Speech Development of a Bilingual Child, Vols. 1–4*. Evanston.
- LEVIN, B, and M. RAPPAPORT-HOVAV. (1995). *Unaccusativity*. Cambridge, Mass.: MIT Press.
- LI, Y. A. (1989). *Order and Constituency in Mandarin Chinese*. Dordrecht: Kluwer.
- LIDDELL, S. (1980). *American Sign Language Syntax*. The Hague: Mouton.
- LIGHTFOOT, D. (1979). *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- (1981). ‘The History of NP-Movement’, in C. Baker and J. McCarthy (eds.), *The Logical Problem of Language Acquisition*. Cambridge, Mass.: MIT Press, pp. 86–119.
- (1989). ‘The Child’s Trigger Experience: Degree-0 Learnability’. *Behavioral and Brain Sciences* 12: 321–34.
- (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, Mass.: MIT Press.
- (1999). *The Development of Language*. Oxford: Blackwell.
- (2002). ‘Myths and the Prehistory of Grammars’. *Journal of Linguistics* 38: 113–136.
- (2006). *How New Languages Emerge*, Cambridge: Cambridge University Press.
- LONGOBARDI, G. (1994). ‘Proper Names and the Theory of N-movement in Syntax and Logical Form’. *Linguistic Inquiry* 25: 609–65.
- (2001). ‘Formal Syntax, Diachronic Minimalism, and Etymology: The History of French *Chez*’. *Linguistic Inquiry* 32: 275–302.
- (2003). ‘Methods in Parametric Linguistics and Cognitive History’. Ms., University of Trieste.
- LOPORCARO, M. (1998). *Sintassi comparata dell’accordo participiale romanzo*. Turin: Rosenberg and Sellier.
- LOS, B. (1998). *Infinitival Complementation in Old and Middle English*. The Hague: Holland Academic Graphics.
- LUMSDEN, J. (1999). ‘Language Acquisition and Creolization’, in M. Degraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass.: MIT Press, pp. 129–58.
- LYELL, C. (1830–4). *Principles of Geology*. London.
- MAC EOIN, G. (1994). ‘Irish’, in M. Ball and J. Fife (eds.), *The Celtic Languages*. London: Routledge, pp. 101–44.
- MACMAHON, A. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- (2005). *Quantitative Methods in Language Comparison, Transactions of the Philological Society Special Issue*, 103.2.

- and R. MACMAHON (2003). 'Finding Families: Quantitative Methods in Language Classification'. *Transactions of the Philological Society* 101: 7–56.
- MAHAJAN, A. (1994). 'Universal Grammar and the Typology of Ergative Languages'. Ms., University of California, Los Angeles.
- MANOLESSOU, I. (2001). 'The Evolution of the Demonstrative System in Greek'. *Journal of Greek Linguistics* 2: 119–48.
- MANZINI, M.-R., and K. WEXLER. (1987). 'Parameters, Binding Theory, and Learnability'. *Linguistic Inquiry* 18: 413–44.
- and L. SAVOIA (2005). *I dialetti italiani e romanci*. Alessandria: Edizioni dell'Orso (3 volumes).
- MARTINET, A. (1955). *Economie des changements phonétiques*. Bern: A. Franck.
- MARTINS, A.-M. (2000). 'Polarity Items in Romance: Underspecification and Lexical Change', in S. Pintzuk, G. Tsoulas, and A. Warner (eds.), *Diachronic Syntax: Models and Mechanisms*. Oxford: Oxford University Press, pp. 191–219.
- MARTIN-JONES, M. (2003). 'Diglossia', in W. Frawley (ed.), *The International Encyclopedia of Linguistics (Second Edition)*, Vol 2. Oxford: Oxford University Press, pp. 435–8.
- MASSAM, D. (2000). 'VSO and VOS: Aspects of Niuean Word Order', in A. Carnie and E. Guilfoyle (eds.), *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press, pp. 97–116.
- (2005). 'Predicate Fronting and Lexical Category in Niuean', in A. Carnie, H. Harley, and S. Dooley (eds.), *Verb First: Studies in Predicate Initial Languages*. Amsterdam: John Benjamins, pp. 227–42.
- and C. SMALLWOOD. (1997). 'Essential Features of Predication in Niuean and English', in K. Kusumoto (ed.), *Proceedings of NELS 27, GLSA*, University of Massachusetts, Amherst. pp. 236–72.
- MATHIEU, E., and I. SITARIDOU. (2005). 'Split *wh*-Constructions in Classical and Modern Greek: A Diachronic Perspective', in M. Batllori, M.-Ll. Hernanz, C. Picallo, and F. Roca (eds.), *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press, pp. 236–50.
- MATTHEWS, P. (2001). *A Short History of Structural Linguistics*. Cambridge: Cambridge University Press.
- MCCLOSKEY, J. (1992). 'Adjunction, selection and embedded verb second'. Ms., University of California at Santa Cruz.
- (2001). 'The Morphosyntax of *WH*-extraction in Irish'. *Journal of Linguistics* 37: 67–100.
- MCCWHORTER, J. (2001). *The Power of Babel (A Natural History of Language)*. London: Heinemann.
- MEILLET, A. 1912. 'L'évolution des formes grammaticales'. Repr. In A. Meillet, 1958, *Linguistique Historique et Linguistique Générale*, Paris: Champion, pp. 130–58.

- MEILLET, A. (1937). *Introduction à l'étude comparative des langues indo-européennes*. Paris: Hachette.
- MENÉNDEZ-PIDAL, R. (1964). *Manual de gramática histórica española*. Madrid: Espasa-Caple.
- MITCHELL, B., and F. ROBINSON. (1992). *A Guide to Old English*. Oxford: Blackwell.
- MOERENHOUT, M., and W. VAN DER WURFF. (2000). 'Remnants of the Old Order: OV in the Paston Letters'. *English Studies* 81: 513–30.
- MORPURGO-DAVIES, A. (1998). *History of Linguistics; Volume IV: Nineteenth-Century Linguistics* (general editor G. Lepschy). London: Longman.
- MUFWENE, S. (1986). 'The Universalist and Substrate Hypotheses Complement One Another', in P. Muysken and N. Smith (eds.), *Substrata vs. Universals in Creole Languages*. Amsterdam: John Benjamins, pp. 129–62.
- (2001). *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- MÜHLHÄUSLER, P. (1986). *Pidgin and Creole Linguistics*. Oxford: Blackwell.
- MÜLLER, G. (2004a). 'Verb-Second as vP-First'. *Journal of Comparative Germanic Linguistics* 7: 139–274.
- (2004b). 'Argument Encoding and the Order of Elementary Operations'. Paper presented at the 27th GLOW Colloquium, Aristotle University of Thessaloniki.
- MUNARO, A. (2005). 'Grammaticalisation, Reanalysis and CP Layering', in M. Batllori, M.-Ll. Hernanz, C. Picallo, and F. Roca (eds.), *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press, pp. 29–47.
- Ó MURCHÚ, M. (1993). 'Aspects of the Societal Status of Modern Irish', in M. Ball and J. Fife (eds.), *The Celtic Languages*. London: Routledge, pp. 471–90.
- MUYSKEN, P. (1988). 'Are Creoles a Special Type of Language?', in F. Newmeyer (ed.), *Linguistics: The Cambridge Survey, Volume 2*. Cambridge: Cambridge University Press, pp. 285–302.
- (2000). *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- NAKLEH, L., D. RINGE, and T. WARNOW. (2005). 'Perfect Phylogenetic Networks: A New Methodology for Reconstructing the Evolutionary History of Natural Languages'. *Language* 81: 382–420.
- , T. WARNOW, D. RINGE, and S. EVANS. (2005). 'A Comparison of Phylogenetic Reconstruction Methods on an Indo-European Dataset', in A. MacMahon (ed.), *Quantitative Methods in Language Comparison, Transactions of the Philological Society Special Issue*, 103: 171–92.
- NEWMAYER, F. (2004). 'Against a Parameter-setting Approach to Language Variation', in P. Pica, J. Rooryk, and J. van Craenenbroek (eds.), *Language Variation Yearbook, volume 4*. Amsterdam: Benjamins, pp. 181–234.
- (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

- NICHOLIS, M. (2004). *On Pro-Drop*. PhD Dissertation, University of Siena.
- NICHOLS, J. (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago.
- NIYOGI, P. (2004). *The Computational Nature of Language Learning and Evolution*. Ms., University of Chicago. Published as P. Niyogi (2006). *The Computational Nature of Language Learning and Evolution*. Cambridge Mass.: MIT Press.
- and R. BERWICK. (1995). 'The Logical Problem of Language Change'. A.I. Memo No. 1516, MIT Artificial Intelligence Laboratory.
- and — (1997). 'A Dynamical Systems Model for Language Change'. *Complex Systems* 11: 161–204.
- NOMURA, T. (1993). 'Joudaigo no *no to ga nitsuite*'. *Kokubo-Kokubun* 62.2: 1–17, 62.3: 30–49.
- NYLANDER, D. (1986). 'Short Note'. *Journal of Pidgin and Creole Languages* 1: 153–158.
- OSGOOD, C., and T. SEBEEK. (1954). 'Psycholinguistics: A survey of Theory and Research Problems'. *Journal of Abnormal and Social Psychology* 49: 1–203.
- OUHALLA, J. (1994). *Introducing Transformational Grammar*. London: Edward Arnold.
- OWEN JONES, R. (1993). 'The Sociolinguistics of Welsh,' in M. Ball and J. Fife (eds.), *The Celtic Languages*. London: Routledge, pp. 536–605.
- PAGLIANO, C. (2003). *L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie: parles-en de ta numération? Impossible*. PhD Dissertation, University of Nice Sophia Antipolis.
- PAUL, H. (1920). *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer.
- PENNER, Z., and T. BADER. (1995). 'Issues in the Syntax of Subordination: A Comparative Study of the Complementizer System in Germanic, Romance and Semitic Languages with Special Reference to Swiss German', in Z. Penner (ed.), *Topics in Swiss German Syntax*. Bern: Lang, pp. 73–290.
- PENNY, R. (1991). *A History of the Spanish Language*. Cambridge: Cambridge University Press.
- PERLMUTTER, D. (1978). 'Impersonal Passives and the Unaccusative Hypothesis'. *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, Ca.: Berkeley Linguistics Society, pp. 157–89.
- PESETSKY, D. (1994). *Zero Syntax*. Cambridge, Mass.: MIT Press.
- PHILLIPS, C. (1995). 'Syntax at Age Two: Cross-Linguistic Differences', in C. Schütze, K. Broihier, and J. Ganger (eds.), *Papers on Language Processing and Acquisition, MIT Working Papers in Linguistics*, 26: 1–58.
- PIERCE, A. (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht: Kluwer.
- PINKER, S., and R. JACKENDOFF. (2005). 'What's Special about the Human Language Faculty?'. *Cognition* 95: 201–36.
- PINTZUK, S. (1991). *Phrase Structure in Competition: Variation and Change in Old English Word Order*. PhD Dissertation, University of Pennsylvania.

- PINTZUK, S. (1993). 'Verb Seconding in Old English: Verb Movement to Infl'. *The Linguistic Review* 10: 5–35.
- (1999). *Phrase Structures in Competition. Variation and Change in Old English Word Order*. New York: Garland.
- (2002). 'Verb-object Order in Old English: Variation as Grammatical Competition,' in D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, pp. 276–99.
- (2003). 'Variationist Approaches to Syntactic Change,' in R. Janda and B. Joseph (eds.) *Handbook of Historical Linguistics*. Oxford: Blackwell, pp. 509–528.
- (2005). 'The Syntax of Objects in Old English', in M. Batllori, M.-Ll. Hernanz, C. Picallo, and F. Roca (eds.), *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press, pp. 251–66.
- and A. KROCH. (1989). 'The Rightward Movement of Complements and Adjuncts in the Old English of *Beowulf*', *Language Variation and Change* 1: 115–43.
- PLATH, S. (1983). *The Journals of Sylvia Plath*. T. Hughes and F. McCollough (eds.). New York: Ballantine.
- PLATZACK, C. (1992). 'Functional Categories in Early Swedish', in J. Meisel (ed.), *The Acquisition of Verb Placement*. Dordrecht: Kluwer, pp. 63–82.
- and A. HOLMBERG (1989). 'The Role of Agr and Finiteness in Germanic VO Languages'. *Working Papers in Scandinavian Syntax* 43: 51–76. Lund: Department of Scandinavian Languages.
- POEPPEL, D., and K. WEXLER. (1993). 'The Full Competence Hypothesis of Clause Structure in Early German'. *Language* 69: 365–424.
- POLETTI, C. (1995). 'The Diachronic Development of Subject Clitics in North-Eastern Italian Dialects', in A. Battye and I. Roberts (eds.), *Clause Structure and Language Change*. Oxford: Oxford University Press, pp. 295–324.
- (2000). *The Higher Functional Field in the Northern Italian Dialects*. Oxford: Oxford University Press.
- POLLOCK, J.-Y. (1986). 'Sur la syntaxe de EN et le paramètre du sujet nul,' In M. Ronat and D. Couquaux (eds.), *La Grammaire Modulaire*. Paris: Les éditions de minuit, 211–46.
- (1989). 'Verb Movement, Universal Grammar, and the Structure of IP'. *Linguistic Inquiry* 20, 365–424.
- POPLACK, S. (1980) ' "Sometimes I'll start a sentence in Spanish *y termino en español*": Toward a typology of code-switching'. *Linguistics* 18: 581–618.
- PRICE, G. (1971). *The French Language: Present and Past*. London: Edward Arnold.
- PULLUM, G. (2003). 'Learnability'. *The Oxford International Encyclopedia of Linguistics*, 431–4. Oxford: Oxford University Press.
- and B. SCHOLZ. (2002). 'Empirical Assessment of Stimulus Poverty Arguments'. *The Linguistic Review* 19: 9–50.

- RACKOWSKI, A., and L. TRAVIS. (2000). 'V-initial Languages: X or XP Movement and Adverbial Placement', in A. Carnie, and E. Guilfoyle (eds.), *The Syntax of Verb-Initial Languages*. Oxford: Oxford University Press, pp. 117–42.
- RADFORD, A. (1990). *Syntactic Theory and the Acquisition of English Syntax*. Oxford: Blackwell.
- (1996). 'Towards a Structure-Building Model of Acquisition,' in H. Clahsen (ed.), *Generative Perspectives on Language Acquisition*. Amsterdam: Benjamins, pp. 43–89.
- (1997). *Syntactic Theory and the Structure of English*. Cambridge: Cambridge University Press.
- (2004). *Minimalist Syntax: Exploring the structure of English*. Cambridge: Cambridge University Press.
- RENZI, L. (1983). 'Fiorentino e italiano: storia dei pronomi personali soggetto', in F. Albano Leoni *et al.* (eds.), *Italia linguistica: idee, storia, struttura*. Bologna: Il Mulino, pp. 223–9.
- RICHARDS, M., and T. BIBERAUER. (2005). 'Explaining EXPL', in M. den Dikken and C. Tortora (eds.), *The Function of Function Words and Functional Categories*. Amsterdam: Benjamins, pp. 115–54.
- RINGE, D., T. WARNOW, and A. TAYLOR (2002). 'Indo-European and Computational Cladistics'. *Transactions of the Philological Society* 100: 59–130.
- RITTER, B. (1991). 'Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew', in S. Rothstein (ed.), *Syntax and Semantics* 26: 37–62. San Diego: Academic Press.
- RIZZI, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
- (1986a). 'Null Objects in Italian and the Theory of *pro*'. *Linguistic Inquiry* 17: 501–57.
- (1986b). 'On the Status of Subject Clitics in Romance', in O. Jaeggli and C. Silva-Corvalán (eds.), *Studies in Romance Syntax*. Dordrecht: Foris, pp. 391–419.
- (1989). 'On the Format for Parameters'. *Brain and Behavioral Sciences* 12: 355–6.
- (1990). *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- (1994). 'Root Null Subjects and Early Null Subjects,' in T. Hoekstra and B. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins, pp. 151–76.
- (1999). 'Broadening the Empirical Basis of Universal Grammar Models: A Commentary', in M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass.: MIT Press, pp. 453–72.
- (2000). 'Relativized Minimality Effects', in M. Baltin and C. Collins (eds.), *Handbook of Syntactic Theory*. Oxford: Blackwell, pp. 89–110.
- (2005). 'On the Grammatical Basis of Language Development: A Case Study', in G. Cinque and R. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*. Oxford: Oxford University Press, pp. 70–109.

- RIZZI, L. and I. ROBERTS. (1989). 'Complex Inversion in French'. *Probus* 1: 1–30.
- ROBERGE, Y. (1990). *The Syntactic Recoverability of Null Arguments*. Montreal: McGill-Queens University Press.
- (1998). 'Les prépositions orphelines dans diverses variétés de français d'Amérique du Nord,' in P. Brasseur (ed.), *Français d'Amérique: variation, créolisation, normalisation*. Centre d'études canadiennes, Université d'Avignon, pp. 49–60.
- (1999). 'Preposition Stranding and *que*-deletion in Varieties of North American French'. *Linguistica Atlantica* 21: 153–68.
- and M.-T. VINET (1989). *La variation dialectale en grammaire universelle*, Montréal: Les Presses de l'Université de Montréal.
- ROBERT (1992). *Dictionnaire Historique de la Langue Française*, sous la direction d'Alain Rey. Paris: Dictionnaires le Robert.
- ROBERTS, I. (1985). 'Agreement Parameters and the Development of English Modal Auxiliaries'. *Natural Language and Linguistic Theory* 3: 21–58.
- (1993a). *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.
- (1993b). 'The Nature of Subject Clitics in Franco-Provençal Valdôtain,' in A. Belletti (ed.), *Dialects of Italy*. Turin: Rosenberg and Sellier, pp. 319–53.
- (1994). 'Two Types of Head Movement in Romance', in N. Hornstein and D. Lightfoot (eds.), *Verb Movement*. Cambridge: Cambridge University Press, pp. 207–42.
- (1995). 'Object Movement and Verb Movement in Early Modern English,' in H. Haider, S. Olsen, and S. Vikner (eds.), *Studies in Comparative Germanic Syntax*. Dordrecht: Kluwer, pp. 269–84.
- (1996). *Comparative Syntax*. London: Edward Arnold.
- (1997). 'Directionality and Word Order Change in the History of English,' in A. van Kemenade and N. Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, pp. 397–426.
- (1998). Review of A. Harris and L. Campbell *Historical Syntax in Cross-Linguistic Perspective*. *Romance Philology* 51: 363–70.
- (1999). 'Verb Movement and Markedness', in M. DeGraff (ed.), *Language Creation and Language Change*. Cambridge, Mass.: MIT Press, 287–328.
- (2001). 'Language Change and Learnability', in S. Bertolo (ed.), *Parametric Linguistics and Learnability*. Cambridge: Cambridge University Press, pp. 81–125.
- (2005). *Principles and Parameters in a VSO Language: A Case Study in Welsh*. Oxford: Oxford University Press.
- and A. HOLMBERG. (2005). 'On the Role of Parameters in Universal Grammar: a Reply to Newmeyer', in H. Broekhuis, N. Corver, M. Everaert, and J. Koster (eds.), *Organising Grammar: A Festschrift for Henk van Riemsdijk*. Berlin: Mouton de Gruyter.
- and M. KATO. (1993). *Viagem Diacrônica pelas Fases do Português Brasileiro: Homagem a Fernando Tarallo*. Campinas: Editora da Unicamp.

- and A. ROUSSOU. (1999). 'A Formal Approach to "Grammaticalisation"'. *Linguistics* 37: 1011–41.
- and — (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.
- ROEPER, T., and B. ROHRBACHER. (2000). 'True Pro-drop in Child English and the Principle of Economy of Projection', in C. Hamann and S. Powers (eds.), *The Acquisition of Scrambling and Cliticisation*. Dordrecht: Kluwer, pp. 351–86.
- RÖGNVALDSSON, E. (1996). 'Word Order Variation in the VP in Old Icelandic'. *Working Papers in Scandinavian Syntax* 58: 55–86.
- ROHLFS, G. (1969). *Grammatica storica della lingua italiana e dei suoi dialetti. Sintassi e formazione delle parole*. Turin: Einaudi.
- ROHRBACHER, B. (1997). *Morphology-Driven Syntax*. Amsterdam: Benjamins.
- ROMAINE, S. (1988). *Pidgin and Creole Languages*. London: Longman.
- ROSS, J. (1967). *Constraints on Variables in Syntax*. PhD Dissertation, MIT.
- ROSSI, M.-A. G. L. (1993). 'Estudo diacrônico sobre as interrogativas do português do Brasil', in I. Roberts and M. Kato (eds.), *Viagem Diacrônica pelas Fases do Português Brasileiro: Homenagem a Fernando Tarallo*. Campinas: Editora da Unicamp, pp. 307–42.
- ROTTET, K. (1993). 'Functional Categories and Verb Raising in Louisiana Creole'. Paper given at the Society of Pidgin and Creole Linguistics, Los Angeles.
- ROUSSOU, A. (1991). 'Nominalized Clauses in the Syntax of Modern Greek'. *UCL Working Papers in Linguistics* 3: 77–100.
- (1994). *The Syntax of Complementisers*. PhD Dissertation, University College London.
- and I. TSIMPLI (2006). 'On (Greek) VSO again!'. *Journal of Linguistics* 42: 317–54.
- ROUVERET, A. (1994). *Syntaxe du gallois: principes généraux et typologie*. Paris: CNRS Editions.
- ROWLANDS, E. (1969). *Teach Yourself Yoruba*. London: English Universities Press.
- RUDIN, C. (1988). 'On Multiple Questions and Multiple WH Fronting'. *Natural Language and Linguistic Theory* 6: 445–502.
- RUSSELL, P. (1995). *An Introduction to the Celtic Languages*. London: Longman.
- SANDLER, W., and D. LILLO-MARTIN. (2001). 'Natural Sign Languages', in M. Aronoff and J. Rees-Miller (eds.), *The Handbook of Linguistics*. Oxford: Blackwell, 563–81.
- SANKOFF, G., and S. LABERGE. (1973). 'On the Acquisition of Native Speakers by a Language'. *Kivung* 6: 32–47.
- SANTORINI, B. (1989). *The Generalization of the Verb-second Constraint in the History of Yiddish*. PhD Dissertation, University of Pennsylvania.
- (1992). 'Variation and Change in Yiddish Subordinate Clause Word Order'. *Natural Language and Linguistic Theory* 10: 595–640.
- (1993). 'The Rate of Phrase Structure Change in the History of Yiddish'. *Language Variation and Change* 5: 257–83.

- SAPIR, E. (1921). *Language*. New York: Harcourt Brace and Co.
- SAUSSURE, F. de (1959). *Course in General Linguistics*. New York: McGraw Hill.
- SAWYER, P. (1971). *The Age of the Vikings*. London: Edward Arnold.
- SCHLEGEL, F. (1808). *Über die Sprache und Weisheit der Indier. Ein Betrag zur Begründung der Alterthumskunde*. Heidelberg: Mohr and Zimmer.
- SCHLEICHER, A. (1850). *Die Sprachen Europas in systematischer Uebersicht* (Linguistische Untersuchungen, II). Bonn: König.
- (1861–2). *Compendium der vergleichenden Grammatik der indogermanischen Sprachen. Kurzer Abriss einer Laut- und Formenlehre der indogermanischen Ursprache, des Altindischen, Altiranischen, Altgriechischen, Altitalischen, Altkeltischen, Altslawischen, Litauischen und Altdeutschen*. Weimar: Böhlau.
- SCHUCHARDT, H. (1979). *The Ethnography of Variation. Selected Writings on Pidgins and Creoles*, T. Markey (ed. and trans.), Introduction by D. Bickerton. Ann Arbor: Karoma.
- (1980). *Pidgin and Creole Languages: Selected Essays*, G. Gilbert (ed. and trans.). Cambridge: Cambridge University Press.
- SCHÜTZE, C. (1997). *INFL in Child and Adult Language: Agreement, Case and Licensing*. PhD Dissertation, MIT.
- SCHWARTZ, B., and R. SPROUSE. (1996). 'L2 Cognitive States and the Full Transfer/ Full Access Model'. *Second Language Research* 12: 40–72.
- and S. VIKNER. (1996). 'The Verb Always Leaves IP in V2 Clauses,' in A. Belletti and L. Rizzi (eds.), *Parameters and Functional Heads*. New York: Oxford University Press, pp. 11–62.
- SENGHAS, A. (1995a). 'The Development of Nicaraguan Sign Language via the Language Acquisition Process', in D. McLaughlin and S. McEwen (eds.), *Proceedings of the 19th Annual Boston University Conference on Language Development*. Somerville, Mass.: Cascadilla Press, pp. 543–52.
- (1995b). *Children's Contribution to the Birth of Nicaraguan Sign Language*. PhD Dissertation, MIT.
- J. KEGL, R. SENGHAS, and M. COPPOLA. (1994). 'Sign Language Emergence and Sign Language Change: Children's Contribution to the Birth of a Language'. Poster presented at the Linguistic Society of America, Boston, Mass.
- SEUREN, P. (1998). *Western Linguistics: An Historical Introduction*. Oxford: Blackwell.
- SHAFFER, R. (1994). *Nonfinite Predicate Initial Constructions in Modern Breton*. PhD Dissertation, University of California, Santa Cruz.
- SHLONSKY, U. (1997). *Clause Structure and Word Order in Hebrew*. Oxford: Oxford University Press.
- SIGURÐSSON, H. (1989). *Verbal Syntax and Case in Icelandic*. PhD Dissertation, University of Lund.
- SIMONE, R. (1998). 'The Early Modern Period', in G. Lepschy (ed.), *History of Linguistics, Volume III: Renaissance and Early Modern Linguistics*. London: Longman, pp. 149–236.

- SIHLER, A. (1995). *A New Comparative Grammar of Greek and Latin*. Oxford: Oxford University Press.
- SIMPSON, A., and Z. WU. (2001). 'IP-Raising, Tone Sandhi and the Creation of Sentence-final Particles'. *Journal of East Asian Linguistics* 11: 67–99.
- SMITH, I. (1978). 'Sri Lanka Creole Portuguese phonology'. *International Journal of Dravidian Linguistics* 7: 248–406.
- (1979a). 'Convergence in South Asia: A Creole Example'. *Lingua* 48: 193–222.
- (1979b). 'Substrata Versus Universals in the Formation of Sri Lanka Portuguese'. *Papers in Pidgin and Creole Linguistics*. Canberra: Pacific Linguistics, no. 2 A-57, pp. 183–200.
- (1984) 'The Development of Morphosyntax in Sri Lanka Portuguese', in M. Sebba and L. Todd (eds.), *Papers from the York Creole Conference, September 24–27, 1983, York Papers in Linguistics* 11, Department of Language, University of York.
- SMITH, N. (2004). *Chomsky: Ideas and Ideals*. Cambridge: Cambridge University Press.
- and I. TEMPLI. (1995). *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.
- SONG, J. J. (2001). *Linguistic Typology: Morphology and Syntax*. London: Longman.
- SPENCER, A. (1991). *Morphological Theory*. Oxford: Blackwell.
- SPORTICHE, D. (1998). 'Subject Clitics in French and Romance, Complex Inversion and Clitic Doubling,' in K. Johnson and I. Roberts (eds.), *Beyond Principles and Parameters: Essays in Memory of Osvaldo Jaeggli*. Dordrecht: Kluwer, pp. 189–222.
- SPOUSE, R., and B. VANCE. (1999). 'An Explanation for the Decline of Null Pronouns in Certain Germanic and Romance Languages', in M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, MA: MIT Press, pp. 257–84.
- STOCKWELL, R. (1977). 'Motivations for Exbraciation in Old English', in C. Li (ed.), *Mechanisms of Syntactic Change*. Austin: University of Texas Press, pp. 291–314.
- and D. MINKOVA (1991). 'Subordination and Word Order Change in the History of English', in D. Kastovsky (ed.), *Historical English Syntax*. Berlin: Mouton de Gruyter, pp. 367–408.
- SYBESMA, R. (1999). *The Mandarin VP*. Dordrecht: Kluwer.
- SYEA, A. (1992). 'Null Subjects in Mauritian Creole and the Pro-Drop Parameter,' in F. Byrne and J. Holm (eds.), *Atlantic Meets Pacific*, pp. 91–104.
- SZEMERÉNYI, O. (1996). *Introduction to Indo-European Linguistics*. Oxford: Oxford University Press.
- TALLERMAN, M. (1987). *Mutation and the Syntactic Structure of Modern Colloquial Welsh*. PhD Dissertation, University of Hull.
- (1996). 'Fronting Constructions in Welsh', in R. Borsley and I. Roberts (eds.), *The Syntax of the Celtic Languages*. Cambridge: Cambridge University Press, pp. 97–124.

- TARALLO, F. (1989). *Fotografias Sociolingüísticas*. Campinas: Editora da Unicamp.
- (1993). 'Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX', in I. Roberts and M. Kato (eds.), *Viagem Diacrônica pelas Fases do Português Brasileiro: Homenagem a Fernando Tarallo*. Campinas: Editora da Unicamp, pp. 69–106.
- TAYLOR, A. (1994). 'The Change from SOV to SVO in Ancient Greek'. *Language Variation and Change* 6: 1–37.
- TAYLOR, D. (1977). *Languages of the West Indies*. Baltimore: Johns Hopkins Press.
- THOMAS, A. (1994). 'English in Wales', in R. Burchfield (ed.), *The Cambridge History of the English Language Volume V: English in Britain and Overseas, Origins and Development*. Cambridge: Cambridge University Press, pp. 94–147.
- THOMASON, S. (2003). 'Contact as a Source of Language Change', in R. Janda and B. Joseph (eds.), *Handbook of Historical Linguistics*. Oxford: Blackwell, pp. 687–712.
- and T. KAUFMAN. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley and Los Angeles: University of California Press.
- THOMPSON, R. (1961). 'A Note on Some Possible Affinities between Creole Dialects of the Old World and Those of the New', in R. Le Page (ed.), *Creole Language Studies II, Proceedings of the Conference on Creole Language Studies, University of the West Indies, Mona*. London: Macmillan.
- THRÁINSSON, H. (2003) 'Syntactic Variation, Historical Development and Minimalism', in R. Hendrick (ed.), *Handbook on Minimalist Syntax*, Blackwell, Oxford, pp. 152–91.
- THURNEYSEN, R. (1892). 'Die Stellung des Verbuns im Altfranzösischen'. *Zeitschrift für Romanische Philologie* 16: 289–371.
- (1946). *A Grammar of Old Irish*. Dublin: School of Celtic Studies, Dublin Institute for Advanced Studies.
- TIMBERLAKE, A. (1977). 'Reanalysis and Actualization in Syntactic Change', in C. Li (ed.), *Mechanisms of Syntactic Change*. Austin: University of Texas Press, pp. 141–77.
- TODD, L. (1974). *Pidgins and Creoles*. London: Routledge.
- TRANEL, B. (1981). *Concreteness in Generative Phonology: Evidence from French*. Los Angeles and Berkeley: University of California Press.
- TRAUGOTT, E. (1992). 'Syntax', in R. Hogg (ed.), *The Cambridge History of the English Language Volume I: The Beginnings to 1066*. Cambridge: Cambridge University Press, pp. 168–298.
- and B. HEINE. (1991). *Approaches to Grammaticalization*. Amsterdam: Benjamins, 2 volumes.
- TRAVIS, L. (1984). *Parameters and Effects of Word Order Variation*. PhD Dissertation, MIT.
- TREMBLAY, M., F. DUPUIS, and M. DUFRESNE. (2005). 'The Reanalysis of the French Prepositional System: A Case of Grammaticalization in Competing Grammars',

- in M. Batllori, M.-Ll. Hernanz, C. Picallo, and F. Roca (eds.), *Grammaticalization and Parametric Variation*. Oxford: Oxford University Press, pp. 109–23.
- TRIPS, C. (2002). *From OV to VO in Early Middle English*. Amsterdam: Benjamins.
- TUTTLE, E. (1986). ‘The Spread of ESSE as a Universal Auxiliary in Central Romance’. *Medioevo romanzo* 11: 229–87.
- VALIAN, V. (1990). ‘Syntactic Subjects in the Early Speech of American and Italian Children’. *Cognition* 40: 21–81.
- VANCE, B. (1988). *Null Subjects and Syntactic Change in Medieval French*. PhD Dissertation, Cornell University.
- (1997). *Syntactic Change in Medieval French: Verb Second and Null Subjects*. Dordrecht: Kluwer.
- VANELLI, L., RENZI, and P. BENINCÀ. (1985). ‘Typologie des pronoms sujets dans les langues romanes’. *Actes du XIIe Congrès des Linguistique et Philologie Romanes*. Aix-en-Provence.
- VENNEMANN, T. (1974). ‘Topics, Subjects, and Word Order: From SXV to SVX via TVX,’ in J. Anderson and C. Jones (eds.), *Historical Linguistics: Proceedings of the First International Congress of Historical Linguistics, Edinburgh, September 1973*, vol. II, Amsterdam: North-Holland, pp. 339–76.
- VIKNER, S. (1995). *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford University Press, Oxford.
- (1997). ‘V-to-I movement and Inflection for Person in all Tenses’, in L. Haegeman (ed.), *The New Comparative Syntax*. London: Longman, pp. 189–213.
- VINCENT, N. (1982). ‘The Development of the Auxiliaries HABERE and ESSE in Romance’, in M. Harris and N. Vincent (eds.), *Studies in the Romance Verb*. London: Croom Helm, pp. 71–96.
- (1988). ‘Latin’, in M. Harris and N. Vincent (eds.), *The Romance Languages*. London: Routledge, pp. 26–78.
- and I. ROBERTS (1999). ‘Remarks on Syntactic Reconstruction’. Talk given at the Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft, University of Constance, February 1999.
- VINET, M.-T. (1984). ‘La syntaxe du québécois et les emprunts de l’anglais’. *Revue de l’association québécoise de linguistique* 3: 221–4.
- VISSER, F. (1963–73). *An Historical Syntax of the English Language*. Leiden: Brill.
- WACKERNAGEL, J. (1892). ‘Über ein Gesetz der Indogermanischen Wortstellung’. *Indogermanische Forschungen* 1: 333–436.
- (1926–8). *Vorlesungen über Syntax*, 2 volumes. Basle: E. Birkhäuser.
- WADDINGTON, C. (1977). *Tools for Thought*. London: Jonathan Cape.
- WANG, Q., D. LILLO-MARTIN, C. BEST, and A. LEVITT. (1992). ‘Null Subject versus Null Object: Some Evidence from the Acquisition of Chinese and English’. *Language Acquisition* 2: 221–54.
- WARNER, A. (1983). Review of Lightfoot (1979). *Journal of Linguistics* 19: 187–209.

- WARNER, A. (1993). *English Auxiliaries: Structure and History*. Cambridge: Cambridge University Press.
- (1997). 'The Structure of Parametric Change, and V Movement in the History of English', in A. van Kemenade and N. Vincent (eds.), *Parameters of Morpho-syntactic Change*. Cambridge: Cambridge University Press, pp. 380–93.
- WATANABE, A. (1996). *Case Absorption and WH-Agreement*. Dordrecht: Kluwer.
- (2001). 'Wh-in-situ Languages', in M. Baltin and C. Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 203–25.
- (2002). 'Loss of Overt Wh-Movement in Old Japanese', in D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, pp. 179–95.
- (2004). 'The Genesis of Negative Concord: Syntax and Morphology of Negative Doubling'. *Linguistic Inquiry* 35: 559–612.
- WATKINS, C. (1976). 'Towards Proto-Indo-European Syntax: Problems and Pseudo-Problems', in S. Seever, C. Walker, and S. Mufwene (eds.), *Papers from the Parasession on Diachronic Syntax*. Chicago: Chicago Linguistic Society, pp. 305–26.
- WEINREICH, U., W. LABOV, and W. HERZOG. (1968). 'Empirical Foundations for a Theory of Language Change', in W. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press, pp. 95–195.
- WEISSENBORN, J. (1992). 'Null Subjects in Early Grammars: Implications for Parameter-setting Theories', in J. Weissenborn, H. Goodluck, and T. Roeper (eds.), *Theoretical Issues in Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum.
- WEXLER, K. (1992). 'Some Issues in the Growth of Control', in R. Larson, S. Iatridou, U. Lahiri, and J. Higginbotham (eds.), *Control and Grammar*. Dordrecht: Kluwer, pp. 253–95.
- (1994). 'Optional Infinitives, Head Movement and the Economy of Derivations', in D. Lightfoot and N. Hornstein (eds.), *Verb Movement*. Cambridge: Cambridge University Press, pp. 305–82.
- (1998). 'Very Early Parameter Setting and the Unique Checking Constraint: A New Explanation of the Optional Infinitive Stage'. *Lingua* 106: 23–79.
- WHEELER, M. (1988). 'Catalan', in M. Harris and N. Vincent (eds.), *The Romance Languages*. London: Routledge, pp. 170–208.
- WHINNOM, K. (1965). 'Contacts de langues et emprunts lexicaux: the origin of the European-based creoles and pidgins'. *Orbis* 14: 509–27.
- WHITE, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- WILLIAMS, A. (2000). 'Null Subjects in Middle English Existentials', in S. Pintzuk, G. Tsoulas, and A. Warner (eds.), *Diachronic Syntax. Models and Mechanisms*. Oxford: Oxford University Press, pp. 164–87.
- WILLIAMS, E. (1994). *Thematic Structure in Syntax*. Cambridge, Mass.: MIT Press.

- WILLIS, D. (1998). *Syntactic Change in Welsh: A Study of the Loss of Verb Second*. Oxford: Clarendon Press.
- (forthcoming). 'A Minimalist Approach to Jespersen's Cycle in Welsh', to appear in S. Anderson and D. Jonas (eds.), *Proceedings of the 8th Diachronic Syntax Conference*.
- WOODCOCK, E. (1959). *A New Latin Syntax*. London: Methuen.
- WU, Z. (2000). *Grammaticalization and the Development of Functional Categories in Chinese*. PhD Dissertation, University of Southern California.
- WUNDT, W. (1880). *Logik. Eine Untersuchung der Prinzipien der Erkenntnis und der Methoden Wissenschaftlicher Forschung*. Stuttgart: Enke.
- WURFF, W. VAN DER (1997). 'Deriving Object-verb Order in Late Middle English'. *Journal of Linguistics* 33: 485–509.
- (1999). 'Objects and Verbs in Modern Icelandic and Fifteenth-century English: A Word Order Parallel and its Causes'. *Lingua* 109: 237–65.
- and T. FOSTER. (1997). 'Object-verb Word Order in 16th Century English: A Study of its Frequency and Status', in R. Hickey and S. Puppel (eds.), *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on his 60th Birthday*, vol. 1. Berlin: Mouton de Gruyter, pp. 439–53.
- WURMBRAND, S. (2001). *Infinitives. Restructuring and Clause Structure*. Amsterdam: Benjamins.
- YANG, C. (2000). *Knowledge and Learning in Natural Language*. PhD Dissertation, MIT.
- (2002). 'Grammar Competition and Language Change', in D. Lightfoot (ed.), *Syntactic Effects of Morphological Change*. Oxford: Oxford University Press, pp. 367–80.
- ZAMPARELLI, R. (1995). *Layers in the Determiner Phrase*. PhD Dissertation, University of Rochester.
- ZEIJLSTRA, H. (2004). *Sentential Negation and Negative Concord*. Utrecht: Netherlands Graduate School of Linguistics.
- ZWART, J.-W. (1997). *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.
- ZWICKY, A. and G. PULLUM. (1983). 'Cliticisation vs. Inflection: English *n't*'. *Language* 59: 502–13.

This page intentionally left blank

Index of Subjects

This index covers all chapters. The further reading sections are briefly indexed, indicated by (fr) following the page numbers. References which also appear in the glossary (pages 445–56) are in **bold**.

- abduction (**abductive change**) 123–7,
226–35, 333–5, *see also*
continuations; drift; reanalysis
- abductive reanalysis 122–41, 208
- Accusative Case (ACC) 155–61,
165–8, 172–3, 250–1, *see also*
parameters A-G: parameter G
(accusative subjects)
- accusative + infinitive
construction 162, 165–8, 170,
172–3, 250–1
- acquisition of language, *see* language
acquisition
- actuation of change 315, 320–1, 333–5,
see also transition of change
- adequacy, in linguistic theory 375
- adpositions 96, *see also* parameters
A-G: F4 (adpositions); prepositions
- Agents 149–51
- Agree** 66–77, 123, 139 n. 8, 140, 142,
155–60, 192–3, 255–6
auxiliary selection 301–2
ergative case marking 306–7
format for parameters A-G 268–72
markedness 276–7
morphological case 248–9
negative concord 80, 149
root (optional) infinitives 218
wh- movement 86
- algorithms** for learning 25 n. 2, 229,
321–2, 386, 388
- ambiguity, *see* P-ambiguity
- analytical economy, *see* majority rule
- antisymmetric approaches to word-
order change 189–97
- approximation approaches to language
acquisition 229, 231–5
- argument structure** 149–61
- asymmetric binary opposition in
parameter values 253–6
- asymmetric c-command 66–72, 189
- attractors 192, 350, *see also* Extended
Projection Principle (EPP)
- auxiliaries 42, *see also* parameters A-G:
F2 (auxiliaries)
auxiliary selection 300–5
English 104–5, 127–9, 352–5
modal 127–9, 134–5, 221–2
Welsh 46
word-order 51–3, 58–60, 94–5, 196
- Bach's generalization 86
- backmutation 374
- bare infinitives 162, 163, 164–5
- behaviourist theories of language
acquisition 18
- bilingualism 324, *see also* creoles;
diglossia (syntactic); pidgins

binding theory 212Bioprogram, *see* Language Bioprogram

Hypothesis

bivalent verbs 433

Blocking Effect 330

borrowing 391, 392, 398, *see also*

language contact

borrowing scale 404–5

Branching Direction Theory

(BDT) 96–9, 101–2, 180

bridge verbs 62

Burzio properties 156–7, 301

c-command 66–72, 159, 166, 189,

241 n. 10

cascades of parametric changes 281–2,

341–2, 351–7

Case (abstract) 155–8, 239, 246–51,

301–2

Accusative Case (ACC) 155–61,

165–8, 172–3, 250–1

Dative Case (DAT) 157, 158, 247–51

Chicken-and-Egg Problem 125–7, 129,

130, 132, 138

Chomskyan syntax, *see* entries in Index
of Names; Minimalist Program;

Universal Grammar (UG)

classification of languages, *see* language
typology**classifiers** (in sign language) 432

clause structure:

clausal complementation 161–75,

250–1

clausal subordination

(hypotaxis) 174

clausal truncation 215–17

clause-chaining (parataxis) 174

clitics:

clitic-climbing 35

clitic-doubling 303 n. 8, 305 n. 9

negation 75

pronouns 239, 352

'special' complement 414

subject 30–2, 37–40, 111–12 (fr),

129–32, 215–17

and Very Early Parameter Setting
(VEPS) 211

clustering effects of parametric

change 313–14

code-switching 316, 326–7

cognitive science, linguistics in 6, 443–4

cognitive verbs 165, *see also* accusative

+ infinitive construction

comparative reconstruction, traditional

(phonological) 358–9

competing grammars 187, 309–14,

319–31, 393

complementation, clausal 161–75,

250–1

complementizer (C) category 50, 51–3,

62, 165 n. 20

complementizer-trace effect 29–31,

34–5

diachronic domain 170–4

from prepositions 163–4

wh-movement 82–3, 85–8

complex inversion construction 130

complexity 234–5, 251–66, *see also*

Inertia Principle; markedness;

opacity; simplicity preference

conservatism of the learner 274–5, *see**also* simplicity preference

Constant Rate Effect

(Kroch's) 309–14, 319–20

contact, language, *see* language contact

continuations (of forms in

reconstructions) 361, 365

correspondence problem 361, 363–7

creoles 237, 396, 406–27, 440–1 (fr), *see**also* entries in Index of Languages;

pidgins

in Deaf community 433–8

- critical-period hypothesis** 209 n. 1, 217, 322, 386, 430, 433–4, 436, 442 (fr)
and second languages 237
- Cross-categorical harmony (CCH)** 96, 179–80, 189
- cues** 133, 242–5, 269–72, 409–10, *see also* triggers
- dative alternation 41 n. 9
- dative case 153–4, 200 (fr), 247–51
- Dative Case (DAT) 157, 158, 247–51
- Deaf communities 293 n. 1, 427–38
- degree-0 hypothesis 183–5, 225 n. 6
- Determiner Phrase (DP) structure 5, 204 (fr)
and reconstruction 369–75
- determinism, weak 231–5
- diachronic changes in parameters A-F
parameter A (null-subject) 33–40
parameter B (V-to-T) 56–8
parameter C (V2) 58–64
parameter D (negative concord) 77–81
parameter E (wh-movement) 90–2
parameter F1–6 (Head-complement) 102–8
- Diachronic Reanalysis (DR) 131–2, *see also* parametric change
- diary drop** (written abbreviated registers) 26, 224
- differentiation, orderly 316–19, 334–5, *see also* grammars in competition
- diffusion, compared to change 297
- diglossia** (syntactic) 316, 323–5
- directional language change, *see* drift
- directionality problem 362, 364
- double-object construction 41 n. 9, 150, 151–2
- double-valence (bivalent) verbs 433
- doublets 330–1, 332–3
- drift 295, 340–57
- dynamical systems** 321–2, 348, 349–50, 438
- echo questions 82–3
- Edge Feature 256 n. 17, 277 n. 24
- elsewhere convention 253
- ergative case marking 306–7
- expressivity 276–7
- Extended Projection Principle (EPP) 54 n.17, 192–5, 235, 256, 271–2, 273–7, 307, 308
- extraposition 181–2, 185–6
- factive** interpretations 168–9, 170, 173
- feature-counting approaches 235, 255–6, 260, 262–4
- features:
Edge Feature 256 n. 17, 277 n. 24
Extended Projection Principle (EPP) 54 n.17, 192–5, 235, 256, 271–2, 273–7, 307, 308
feature-counting approaches 235, 255–6, 260, 262–4
formal features 66, 235, 245, 255, 268–72, 299, 351, *see also* Probes and Goals
 φ -features 68, 155–6, 158–9, 192, 271
Negation features 68–81
Operator feature 72, 79, 80, 177
Person, Number and Gender, *see* φ -features
- formal** operations:
Agree, *see* Agree
Merge, *see* Merge
Move, *see* Move
- formal** optionality 238–9, 259–61, 297, 305–9, 331–3
- free inversion 28, 29–32, 34, 412
- fronting (in Welsh English) 400–1
- Full Access (no Transfer) 387
- Full Transfer Full Access 387

- generalization of the input (in language acquisition) 240, 275, 276–7
- Generative grammar** 113 (fr), 155–7, 444, *see also* Chomsky, N. (Names index); Universal Grammar (UG)
- genetics, *see* innateness hypothesis
- goals, probes and 68–77, 86, 155–6, 192, 195, 235, 308
- gradualness 293–314, 334–5, *see also* drift
- grammar, definition of 12–13
- grammars in competition 187, 309–14, 319–31, 393
- grammaticalization** 141–9, 200–1 (fr), 294, 347
- in sign languages 432
- Greenberg's Universal 41, 179
- Head-Movement Constraint 53
- hierarchies:
- of markedness 255–6
 - of parameters 277–82, *see also* cascades of parametric changes
- Holmberg's generalization 55, 57–8
- homesigns** 429–30, 432–3
- hypotaxis 174
- immature grammars 210, 216–18, *see also* maturation approach to UG
- imperfect learning 385–9, 391, 393, 420, 427
- implementation of change, *see* transition of change
- implicational sequences of language changes 347
- implicational universals** 23, 86, 94–102, 253, 280
- Inertia Principle** 199 (fr), 227, 231–2, 233, 248, 350, 356, 365, 371, 439
- infinitives:
- accusative + infinitive construction 162, 165–8, 170, 172–3, 250–1
 - bare 162, 163, 164–5
 - prepositional infinitives 163–4, 171, 173
 - root (optional) infinitives 209, 214–20
- innateness hypothesis** 14–19, 20, 209–10, 437
- interlanguages** 237, 384–9
- intermediate grammars:
- first-language acquisition 209–26
 - second-language acquisition 385–6
- irrealis morphology 87
- islands** 88–9
- Jespersen's Cycle 81, 142–3, 226 n. 7, 318
- Kitsch program 371–2
- language acquisition 14–19, 207–83, 285–6 (fr), 437
- first-language acquisition (L1) 23–4, 209–30, 231–5
 - generalization of the input 240, 275, 276–7
 - gradualness 314 n. 2
 - and hierarchical parameters 280
 - imperfect learning 385–9, 391, 393, 420, 427
 - intermediate grammars 209–26, 385–6
- learnability** 124, 208, 226, 228–9, 284–5 (fr), 328–31
- learning algorithms** 25 n. 2, 229, 321–2, 386, 388
- and markedness 253, 261–4
 - and reanalysis 123–7, 132–4, 140

- second-language acquisition
 (L2A) 126, 383–442
 Subset Principle 256–61
 trans-generational transmission of
 grammar 124–5, 333–4, 340, 341,
 408
- Language Bioprogram Hypothesis 383,
 407–19, 417, 425, 436
- language contact 126, 389–406
 creoles, *see* creoles
 direct 236–42, 390–1, 406
 as cause of parameter-
 resetting 236–42
 indirect 390–1, 398, 399–404
 in reconstruction 371
- language creation (in
 Nicaragua) 427–38, *see also* creoles
- language faculty** 13–15, 209–10, *see*
also Universal Grammar (UG)
- language families 357, *see also* Index of
 Languages
- language pathology 13 n. 1
- language typology** 20, 23–4, 54–5, 64,
 86, 93–102, 117 (fr)
 approaches to drift 342–5
 Branching Direction Theory
 (BDT) 96–9, 101–2, 180
 hierarchical parameters 279–80
 parametric comparison 375
 similarities across creoles 417
 typological drift 282
 typologically-favoured
 borrowing 405
 and word-order change 176–80
- Law of the Excluded Middle 296
- learnability** 124, 208, 226, 228–9, 284–5
 (fr), 328–31
- learning algorithms** 25 n. 2, 229, 321–2,
 386, 388
- learning path 280
- lexical diffusion 297–300, 351
- Linear Correspondence Axiom
 (LCA) 189–97
- linguistic theory, goals of 12
- logical problem of language change** 208,
 230–5, *see also* Regress Problem
- logistic, the (mathematical
 function) 309–14
- majority rule 359, 364
- markedness** 251–66, 284–5 (fr)
 and cascading parametric
 change 342
 change from unmarked to
 marked 275–7
 compared to Darwinian
 fitness 320–1
 and creoles 396, 417–19, 425
 and drift 349–50
 and grammars in competition 328–9
 and parameters 194, 242–3, 253–6,
 272–5, 435–6
 and reconstruction 367–8
 reversal 267
- massive movement 195–7, 248–9
- maturation approach to UG 217–18,
 314 n. 15, 322
- Merge** 5, 22, 41, 123
 as invariant 193
 and the Linear Correspondence
 Axiom (LCA) 190–1
 second Merge 43, 192
- meta-parameters 194
- microparametric change 300–5
- Minimalist Program** 4, 7, 13n. 1,
 217 n. 13, 254 n. 15, 269 n. 20,
 271, 299, 307–8, 331, 375
- morphology 12, 135–9, 381 (fr)
 Blocking Effect 330
 ergative case marking 306–7

- morphology (*cont.*)
 grammaticalization 141–9
 irrealis 87
 loss of case systems 154, 173, 179,
 247–51
 markedness 261–4
 morphologically-driven parameter
 resetting 245–51
 periphrastic *do* 310–12
 in Primary Linguistic Data
 (PLD) 270
 and second languages 388, 389
 of sign languages 431
 unidirectional change 342
- Move** 29, 123, 139 n. 8, 140, 142, 193,
 218, 255–6
- Natural Serialisation Principle
 (NSP) 343
- negation 115 (fr)
 in creoles 416
 Negation features 68–81
 negative concord 64–81, 149,
 317–19, *see also* parameters A-G:
 D (negative concord)
- negative evidence** 15, 17, 256–7,
 259, 260
 indirect negative evidence 261–4
- networks, of parameters 277–82
- Nicaragua, language creation
 in 427–38
- null-subjects 62–4, 111 (fr)
 early null-subjects 219–24, 261
 expletive 248
 loss of 335–9
 null-subject parameter, *see*
 parameters A-G: A (null-subject)
- object control 164
object shift 55, 57–8
- Occam's razor 127
- opacity 131, 132, 232–5, 236–51,
 251–66, *see also* complexity;
 Transparency Principle
- Operands, *see* Operator feature
- operations:
 Agree, *see* Agree
 Merge, *see* Merge
 Move, *see* Move
- Operator feature 72, 79, 80, 177
- optional infinitives, *see* root (optional)
 infinitives
- optionality:
 formal optionality 238–9, 259–61,
 297, 305–9, 331–3
 true optionality 308–9, 331
- orderly differentiation 316–19, 334–5,
 see also grammars in competition
- P-ambiguity:
 and language acquisition
 (child) 232–4, 235, 236–51, 259,
 270
 and language acquisition
 (creoles) 417–18, 419, 426–7, 434,
 439
 and reanalysis 133–9, 149, 439
- P-expression:
 and language acquisition (child) 228,
 232–5, 243–4, 270
 and reanalysis 133–9
- pair-list readings 16–17, 70 n. 24, 84
- parallel development, of languages 357,
 374
- parameter setting:
 and change 266–83
 in language acquisition 116–17 (fr),
 209–26
 missetting 221, 222–4
 resetting 236–51

parameters:

- A-G, *see* parameters A-G
- Baker's 'periodic table' 278–9
- change, *see* parametric change
- comparison 368–75
- default values 269, 269–72, 270, 418
- discrete nature of 295–6
- format for 267–72
- hierarchies of 277–82, 405
- intrinsic ordering of 277–82
- in language acquisition 116–17 (fr),
209–26
- missetting 221, 222–4
- networks of 277–82
- parameter grid 369–70
- parameter interaction 307
- properties 270
- quantity-sensitive stress-assignment
(Dresher) 242–3
- reason for existence 271
- relations between parameters 270
- resetting 236–51

parameters A-G:

- A (null-subject) 24–40, 211, 212,
217–18, 219–24, 257–9, 261, 267–8,
271, 286 (fr), 366, 412–14
- B (V-to-T) 41–8, 45, 54, 56–8, 57, 64,
127–9, 134, 211, 212, 218, 267–8,
271, 309–14
- C (V2) 48–55, 50, 58–64, 61, 64, 211,
212, 267–8, 271, 366–7
- D (negative concord) 64–81, 75,
80–1, 81, 90–2, 149, 211, 212,
267–8, 271
- E (wh-movement) 81–92, 83, 211,
267–8, 272
- F1–6 (Head-complement) 92–108,
211, 212, 267–8, 272, 273–5
- F1 (OV/VO order) 94, 175–98
- F2 (auxiliary order) 95, 107

- F3 (complement of V/T order) 96,
104–8, 175–98, 181, 194
 - F4 (adpositions' objects order) 96
 - F5 (structural complement of V/T/P
overt order) 96
 - F6 (complement of H overt order) 96,
104–8, 177, 178, 180, 181, 194
 - G (accusative subjects) 172, 172–3,
250, 267–8, 272
- parametric change, *see also* diachronic
changes in parameters A-F;
language acquisition; reanalysis
- cascades of 281–2, 341–2, 351–7
 - clustering effects of parametric
change 313–14
 - and drift 345–51
 - gradualness 293–314
 - and hierarchical parameters 280–2
 - linked to Diachronic Reanalysis 131
 - microparametric change 300–5
 - parameter-setting and change 266–82
 - in the principles-and-parameters
approach 109
 - resetting 236–51
- parametric comparison 368–75
- parametric indices 366–7
- perception verbs 165 n. 20
- φ -features 68, 155–6, 158–9, 192, 271
- phonology 12, 288–9 (fr), 381 (fr)
- markedness 251–3
 - orderly differentiation 316–19
 - and reanalysis 129–32
 - reconstruction 358–9
 - reductive phonological change 179
- phylogenetic approach to
reconstruction 367–75
- pidgins** 237, 406, 440–1 (fr), *see also*
creoles
- in Deaf community 429, 430
 - as PLD for creoles 408

pied-piping 84–5, 195–6, 238–9, 241,
260, 269 n. 20

optional 308–9, 332–3

polarity items 72–5, 79

'pool of variants' problem 362, 367–8

postpositions, *see* adpositions;

prepositions

poverty-of-the-stimulus argument 6,
14–19, 140, 193, 212, 232, 437, 443,
see also primary linguistic data

Prague School 251, 253 n. 14

prepositions:

grammaticalization 163–4

Preposition-stranding 84, 238–42,
260, 406, 416

prepositional complementizers 172

prepositional infinitives 163–4, 171,
173

primary linguistic data (PLD), *see also*

triggers; cues

abduction 334

contact 388–9, 390–1, 392

dynamical systems 321–3, 438

markedness 329, 332–3

and the maturation approach 217

opacity 236–51

orderly differentiation 334

P-ambiguity 236–51, 434, 435, 439

and parameter setting 210–12, 228,
230–4

pidgins and creoles 384, 433–4, 439

and poverty-of the stimulus

argument 14–19

trans-generational transmission of
grammar 388–9

principles-and-parameters

approach 21–4, 284 (fr), 444

and creoles 417

language typology 95

and parametric change 109

uniformitarian hypothesis 264–6

Probably Approximately Correct (PAC)

algorithms 229, *see also* learning
algorithms

Probes and Goals 68–77, 86, 155–6,
192, 195, 235, 308

pronouns:

anaphoric 212

overt expletive 221–2

strong 39–40

proto-grammar 360, 365

psychological (psych) verbs 151, 153–5,
159, 160, 200 (fr), 298

quantified expressions, in first-language
acquisition 212

quantity-sensitive stress-assignment
(QS and QI) 242–3

reanalysis 122–41, 198–9 (fr)

and argument structure 152–61

Chicken-and-Egg Problem 125–7,
129, 130, 132, 138, *see also* Regress
Problem

and clausal complementation 173–4

in contact situation 241–2

Diachronic Reanalysis (DR) 131–2

and grammaticalization 141–9

and reconstruction 360–1

recipient passives 152–61, 159, 246–51

reconstruction 357–77

recursivity 12

reflexes, *see* continuations (of forms in
reconstructions)

registers 26, 224, 239, 324–5

Regress Problem 125–7, 129, 130, 132,
138, 208, 230–1, 245, 329, 333, *see also*
logical problem of language change

and contact 405

and second languages 385

relations:

Agree, *see* Agree

relexification 383, 419–25

rightmost-branch problem 190–1

root (optional) infinitives 209, 214–20

semantic roles, *see* thematic roles

sign language:

general 428

homesigns 429–30

in Nicaragua 429–38

simplicity preference 131–2, 134,

231–5, 272, 274

social stratification 316–19

sociolinguistic factors 296–7, 313,

315–40, 381–2 (fr)

speech communities 293 n. 1, 313, 334

diglossia (syntactic) 316, 323–5

spread of syntactic change 315–40

steady states (languages converging

towards) 265–6

stress systems 242–3

stylistic inversion 36 n. 7, *see also* free

inversion

subject-control verbs 164–5

subjunctives 162–4, 173

Subset Principle 256–61, 275, 284 (fr)

substantive universals 262–4, *see also*

feature-counting approaches

substrata 389–406, 407, 434, 436

substratum/ relexification

hypothesis 419–25

successive-cyclic movement 89

superset trap 257, *see also* Subset

Principle

superstrata (lexifier languages) 419–25

terminal nodes, *see* Linear

Correspondence Axiom (LCA)

thematic roles 149–52, 153–8, 181

TMA particles (tense/mood/aspect), in
creoles 414–16, 421

topic-drop 223

topicalization 49, 50

trans-generational transmission of

grammar 124–5, 129, 333–4, 340,

341, 408

transition (implementation) of

change 315–40

transitive-expletive constructions 55–6,

58, 187 n. 32

Transparency Principle 127–9, 181,

346 n. 30

triggers 183, 193–4, 230, 232, 236–51,

see also cues

impoverished 408

true optionality 308–9, 331

typology of languages, *see* language

typology

unbounded dependencies 88–9

unidirectional change 344–5

uniformitarian hypothesis 6, 174,

264–6, 349

Universal Grammar (UG) 7, 71, 124,

210, 264–6, 270–1, 330, 347,

349–50, 385, 386–9, 443, *see also*

Minimalist Program

basics of 12–19

and creoles 407, 408–9, 416

and language typology 95

and the uniformitarian hypothesis 174

V-to-T movement 199–200 (fr),

218–19, 243, 245–51, *see also*

parameters A-G: B (V-to-T)

V2 48–56, 56–64, 60–1, 93, 215–17, 321,

see also parameters A-G: C (V2)

verb-raising 182–3, 197 n. 34

verb serialization 422, 432

verbs:

- bivalent 62, 433
- cognitive 165, *see also* accusative + infinitive construction
- perception 165 n. 20
- psychological (psych) verbs 151, 153–5, 159, 160, 200 (fr), 298
- subject-control 164–5

Very Early Parameter Setting

(VEPS) 211, 280, 425

weak determinism 231–5

wh-in-situ languages 83, 86, 89

wh-movement 81–92,

115–16 (fr), 238, 332–3,

see also parameters

A-G: E (*wh*-movement)

in creoles 416

reconstruction 367

word-order, *see also* parameters A-G

antisymmetric approaches

189–97

change 175–98, 201 (fr), 391–9

markedness 273–5

X'-theory 48, 180–1

Index of Names

The index covers all of the main chapters but not the further reading sections. Only significant mentions of an author or lengthy quotations are indexed.

- D'Alessandro, R. 307
Allen, C. 152–61, 247, 298
Andersen, H. 123, 124
Anderson, S. 137
Avrutin, S. 215
- Bach, E. 86
Baker, M. 277–9, 281, 422
Barbosa, P. 336–8
Bartsch, R. 177
Battistella, E. 251, 253, 259
Bertolo, S. 228
Berwick, R. 256–61, 321
Biberauer, T. 308–9, 332–3
Bickerton, D. 384, 407–19, 425
Bobaljik, J. 56
Bolinger, D. 146–7, 148
Borer, H. 32, 299
Brandi, L. 30–1
Burzio, L. 156–7, 301
- Campbell, L. 122–3, 126–7,
132, 140, 174–5, 294, 325 n. 19,
358, 361
Cardinaletti, A. 26, 27 n. 3, 37, 39, 40
Cecchetto, C. 167–8
Cheng, L. 85–6
Chomsky, N. 4–8, 12–14, 18, 19, 43,
66–8, 71, 89, 139 n. 8, 140, 166,
178, 190–1, 192, 210, 234–5,
248 n.13, 252–3, 254 n. 15, 255,
261–4, 276, 277 n. 24, 299, 308, 320
Cinque, G. 252, 262–4
Clark, R. 25 n. 2, 133, 225 n. 6, 230, 233
Collins, C. 157
Comrie, B. 179
Coppola, M. 428, *see also* KSC
Cordin, P. 30–1
Croft, W. 125 n. 2, 264–6, 342, 343–5,
346
- Darwin, C. 294
DeGraff, M. 384, 398 n. 5, 415, 419,
424–5, 426–7
Denison, D. 296–7
Déprez, V. 70, 145–6, 149 n. 12
Dresher, E. 133, 242–5, 269–72, 280
Dryer, M. 95 n. 30, 96–102, 180
Duarte, E. 336–8
- Ellegård, A. 310–11
Emonds, J. 40, 42
Ernout, A. 162 n. 18, 169–70, 173
- Faarlund, J.-T. 221
Fischer, O. 186, 187, 188, 193
Fitch, W. 140
Fortson, B. 107

- Foulet, L. 77–9, 146–7, 148
Fowler, J. 316 n. 17
Fuß, E. 187 n. 31, 327
- GGL (Gianollo, Guardiano, and Longobardi) 368–75
Giannakidou, A. 74, 76
Gianollo, C. *see* GGL
Gray, D. 135–6
Greenberg, J. 94, 176, 179, 188, 343
Guardiano, C. *see* GGL
Guasti, M.-T. 18, 19, 165 n. 20, 208, 211, 211–12, 214, 217, 219, 220, 256–7, 314 n. 15, 329 n. 24
- Haegeman, L. 26
Hale, M. 296, 297, 360
Halle, M. 234–5, 252–3, 255, 262
Harris, A. 106, 122–3, 126–7, 132, 140, 145 n. 9, 168, 174–5, 294, 325 n. 19, 358, 361
Haspelmath, M. 294
Hauser, M. 140
Hawkins, J. 95, 106, 177 n. 28, 178, 179–80, 194
Héroard, J. 226 n. 7
Herzog, W. *see* WLH
Hinterhölzl, R. 195
Hockett, C. 296
Hoeksema, J. 72 n. 26
Holm, J. 406–7
Holmberg, A. 55, 281 n. 26, 339
Hornstein, N. 240
Huang, J. 81, 89
Hyams, N. 211, 221–3, 261
- Ingham, R. 187 n. 32
- Jakobson, R. 251, 253
Jespersen, O. 81, 354, 391 n. 2
- Jonas, D. 56
Jones, M. 165
- Kato, M. 336–8
Kaufman, T. 396–7, 400, 404–5
Kayne, R. 32, 33, 36 n. 7, 37, 43, 146, 164, 166, 172, 189, 190, 194, 238–9, 300
Keenan, E. 227, 231
Kegl, J. 428, 435–6, 437, 438, *see also* KSC
van Kemenade, A. 60–1, 180–8, 189, 352
King, R. 238–42
Kiparsky, P. 108, 192, 193, 193–4, 195, 275–7
Kroch, A. 125, 126, 187–8, 234 n. 8, 293–4, 294–5, 296, 309–14, 319–20, 323, 328, 330–1, 389
KSC (Kegl, Senghas, and Coppola) 428, 429–38
Kühner, R. 169
- Labov, W. 65 n. 22, 292, 293, 294 n. 3, 316–19, *see also* WLH
Lakoff, R. 345–6
Lasnik, H. 261
Lass, R. 350 n. 32, 360
Ledgeway, A. 303–4
Lehmann, W. 107–8, 176–7, 180, 188, 343
Lightfoot, D. 7, 125, 126, 127–9, 132, 133, 178, 183–7, 188, 189, 225 n. 6, 242–5, 295, 345–7, 350, 356, 358, 360, 360–1, 362, 363 n. 40
Lillo-Martin, D. 428
Longobardi, G. 146, 227, 232, 282, 368–75, 375
- McCloskey, J. 402 n. 6
MacMahon, A. 426
Martins, A.-M. 80–1

- Massam, D. 279 n. 25
 Mathieu, E. 92
 Maupas, C. 35
 Meillet, A. 140
 Morpugo-Davies, A. 340–1
 Mufwene, S. 425–6
 Müller, G. 52 n. 15, 305 n. 10, 306–7
 Muysken, P. 327, 416

 Newmeyer, F. 281 n. 26
 Nichols, J. 265–6, 276
 Niyogi, P. 15, 25 n. 2, 208, 229, 321–4,
 328, 350–1

 Oniga, R. 167–8

 Paul, H. 125, 315
 Pesetsky, D. 151 n. 14
 Phillips, C. 218
 Pierce, C.S. 123–4
 Pintzuk, S. 105, 187, 188, 319, 325–7,
 330
 Polleto, C. 39
 Pollock, J.-Y. 36 n. 7, 40, 42
 Pullum, G. 14–15, 19, 354 n. 35

 Radford, A. 16
 Richards, M. 308–9, 332–3
 Ringe, D. 372–3, 374
 Rizzi, L. 26–7, 29–30, 31–2, 43,
 215–17
 Roberts, I. 25 n. 2, 131–2, 132, 133,
 137, 146, 225 n. 6, 227, 230, 231,
 233, 234–5, 255–6, 281 n. 26, 307,
 308–9, 347, 349, 352–3, 358, 362,
 419, 427
 Rohlfs, G. 170, 304–5
 Ross, J. 88
 Roussou, A. 27 n. 3, 146, 227, 230, 231,
 234–5, 255–6, 347

 Sandler, W. 428
 Sapir, E. 341, 348, 352, 356
 Sawyer, P. 396 n. 3
 Schlegel, F. 174
 Schleicher, A. 341
 Scholz, B. 14–15, 19, 354 n. 35
 Schuchardt, H. 406, 420
 Senghas, A. 428, *see also* KSC
 Seuren, P. 4 n. 4
 Sigurðsson, H. 158 n. 17
 Sihler, A. 171 n. 23
 Sitaridou, I. 92
 Song, J. J. 178
 Starke, M. 37, 40
 Stegmann, C. 169
 Stockwell, R. 183

 Taylor, A. 107, 187–8
 see also Ringe, D.
 Thomas, F. 162 n. 18, 169–70, 173,
 400–1, 403
 Thomason, S. 389, 396–7, 400,
 404–5
 Thráinsson, H. 137, 138, 139
 Thurneysen, R. 62
 Timberlake, A. 126–7, 132
 Trips, C. 187 n. 31, 327, 391–2

 Vance, B. 34 n. 6, 61
 Vennemann, T. 177–9, 179–80, 188,
 294, 300, 343
 Vikner, S. 136, 137
 Vincent, N. 162–3, 357 n. 37, 362

 Warner, A. 56–7, 138
 Warnow, T. *see* Ringe, D.
 Watanabe, A. 69, 72, 86–7, 90–1
 Watkins, C. 361–2
 Weinberg, A. 241
 Weinreich, U. *see* WLH

Wexler, K. 211, 215–17

White, L. 387, 388

Willis, D. 64 n. 21, 313

WLH (Weinreich, Labov
and Herzog) 315–16, 330,
333–5

Woodcock, E. 169

Wundt, W. 4 n. 4

Wurmbrand, S. 165

Yang, C. 321

Zeijlstra, H. 69, 70 n. 24, 71, 75 n. 27, 76

Zwicky, A. 354 n. 35

Index of Languages

The index covers all chapters. The further reading sections are briefly indexed, indicated by (fr) following the page numbers.

- Abruzesse (dialect of Italian) 303, 307
African-American Vernacular English (AAVE) 317–19, 404
Algerian French 38, 39
American Sign Language (ASL) 431
Amerindian languages 86
Amharic 343–5
Atlanta (variant of English) 318–19
Atlantic creoles 409, 420
Australian languages 86
- Baltic languages 179, 373
Balto-Slavonic languages 373
Bambara 421–2
Bantu languages 23, 409
Basque 94, 96, 306–7
Berbice Dutch (creole) 424
Brazilian Portuguese (BP) 91, 335–9
Breton 54
- Calabrian (dialect of Italian) 170
Canadian French 238–42
Cape Verdean (creole) 412, 413, 414
Catalan 80–1
Celtic languages 45, 54, 107, 373, 441 (fr), *see also* individual languages
Chinese 83, 85–6, 89, 223–4, 366 n. 43
creoles, *see also* individual superstrate languages e.g. English, Dutch
Atlantic creoles 409, 420
Berbice Dutch 424
Cape Verdean 412, 413, 414
Crioulo 416
Haitian Creole 410, 413, 414, 415, 423, 425
Indo-Portuguese creoles 411 n. 11
Jamaican 413
Kituba 409
Kriyol 412, 413
Mauritian Creole 410, 413–14, 415
Ndjuka 422
Papiamentu 412, 413, 414, 415
Principe 423
Réunionnais 'semi-creole' 410 n. 10
Saramaccan 412, 413
Trinidad Creole 410
Crioulo (creole) 416
- Danish:
diachronic change in 55–6, 62–3, 138
synchronic variation in 45, 50, 54
- Dutch:
-based creoles 411, 413
diachronic change in 179
synchronic variation in 50, 54, 86, 94, 96, 97, 182, 184–5, 214
- Edo 279, 422
- English:
-based creoles 280, 413, 422
diachronic change in 56–8, 58–61, 64, 102–5, 127–9, 134–7, 138–9,

English (*cont.*)

142–9, 152–61, 175–98, 199–200
(fr), 200 (fr), 201 (fr), 221, 246–51,
280, 298, 302, 310–12, 341, 343,
351–7, 352, 352–3, 391–9

history of 1–3, 117–18 (fr), 441–2 (fr)

parameter settings in 219, 260, 280

synchronic variation in 19–22, 25–7,
29, 41, 42, 45, 50–4, 54, 65, 71–2,
75, 83, 84, 84–5, 92–4, 96, 135–6,
150, 172, 214, 220, 238–9, 240–2,
260, 325, 326–7, 332–3

variants of 65 n. 22, 135–6, 155,
317–19, 318–19, 400–4, 441 (fr)

Ethiopian Semitic languages 343–5

Ewe 421

Faroese 139

Finnish 108, 339

Flemish 182

Florentine (dialect of Italian) 30–1,
37–8

Fongbe 421, 423

Franco-Provençal Valdôtain 129–32

French:

-based creoles 410, 413, 415

diachronic change in 33–40, 61–2,
64, 77–80, 91, 129–32, 142–9, 163,
164, 220, 222, 226 n. 7, 321, 323

history of 110 (fr), 400

phonology 205 (fr)

synchronic variation in 25–6, 28, 29,
32, 41–2, 44–5, 47–8, 50–3, 54,
65–71, 75–6, 85, 87, 172, 211, 214,
238–9, 239, 260, 323

variants of 38, 39, 129–32, 238–42

Friulian (dialect of Italian) 39

Gbe languages 421–2

Ge'ez 343–5

German:

-based creoles 411

diachronic change in 59, 103–5

synchronic variation in 19–22, 26–7,
48–51, 54, 75, 76, 83, 86, 92–4,
95 n. 30, 96, 97, 184–5, 214, 219

variants of 182, 340

Germanic languages 54, 62, 107,

118–19 (fr), 175, 177, 182, 195–6,

202–3 (fr), 221, 307, 308 n. 12,

351–7, 371–2, 373, 394, *see also*

individual languages; Proto-

Germanic; Proto-West-Germanic

Graeco-Armenian languages 373

Greek:

diachronic change in 91–2, 107, 175,
179, 264 n. 19, 309

synchronic variation in 23, 25, 28,
29, 32, 45, 54, 371–2

Guarani 94

Haitian Creole 410, 413, 414, 415,
423, 425

Harari 343–5

Hebrew 46, 264 n. 19, 339

Hiberno-English 400–4, 441 (fr)

Hittite 361

Icelandic, *see also* Old Norse

diachronic change in 107, 179, 192,
239, 355–6, 394, 399

synchronic variation in 45, 50, 54,
158 n. 17, 238, 248 n. 13, 339

Idioma de Señas Nicaragüense
(ISN) 430–8

Indo-European 264 n. 19, 302 n. 5, 343,
346, 377–80 (fr), *see also* individual
languages; language families

reconstruction 107–8, 221, 358, 361,
364–5, 366–7, 372–3

- Indo-Portuguese creoles 411 n. 11
- Iranian languages 343, 344 n. 29
 head-complement order 97
- Irish 54, 87, 107, 400–4
- Italian:
 diachronic change in 221–2, 300–5
 synchronic variation in 25, 27, 28,
 29, 30–1, 32, 45, 50, 54, 75–6, 83,
 163, 164, 170, 280, 300–5
 variants of 30–1, 32, 37–9, 63, 170,
 303 n. 8, 303–5, 306–7, 381 (fr)
- Italo-Celtic languages 373
- Jamaican (creole) 413
- Japanese 23, 25, 83, 90–1, 93, 94, 96, 223
- Kentish (dialect of English) 155
- Kituba (Bantu creole) 409
- Kriyol (creole) 412, 413
- Kronoby (dialect of Swedish) 136
- Kwa languages 420–1, 424
- Latin 23, 203–4 (fr)
 diachronic change in 105–7, 162–75,
 179, 250–1, 264 n. 19
 reconstruction 358, 362, 367–8
 synchronic variation in 86, 97, 371–2
- Lenguaje de Señas Nicaragüense
 (LSN) 430–8
- Malayalam 99–101, 273–4
- Marathi 339
- Mauritian Creole 410, 413–14, 415
- Ndjuka (creole) 422
- Neapolitan (dialect of Italian) 303–4
- New York (variant of English) 316–17,
 318–19
- Nicaraguan sign language 429–38,
 442 (fr)
- Norwegian 54, 136
- Occitan 129–32
- Old Norse 391–9
- Palaun 87
- Papiamentu (creole) 412, 413, 414, 415
- Pidgin de Señas Nicaragüense, El
 (PSN) 430
- Portuguese, Brazilian 91, 335–9
- Portuguese, European (EP) 302 n. 6,
 335–6
 -based creoles 412
- Prince Edward Island (PIE)
 French 238–42
- Principe (creole) 423
- Procidano (dialect of Italian) 303 n. 8
- Proto-Germanic 221
- Proto-Indo-European 107–8, 221
- Proto-Niger-Congo 179
- Proto-West-Germanic 219
- Réunionnais 'semi-creole' 410 n. 10
- Romance languages, *see also* individual
 languages
 -based creoles 280, 409, 414
 diachronic change in 105–7, 162–75,
 179, 250–1, 357
 reconstruction 358, 362, 367–8
 synchronic variation in 23, 54,
 80–1, 94, 96, 145, 203–4 (fr), 357,
 371–2, 414
- Rumanian 76, 170 n. 22, 302 n. 6
- Russian 85, 215, 308–9, 331, 332–3
- Sabir 409 n. 8
- Salentino (dialect of Italian) 170,
 170 n. 22
- Saramaccan (creole) 412, 413
satem group 373
- Scandinavian languages 55–6, 64, 84,
 93, 238, 392, *see also* individual
 languages

- Scots 135 n. 7, 137
- Semitic languages 343–5, 371–2
- Sicilian (dialect of Italian) 170
- sign languages in Nicaragua 429–38,
442 (fr)
- Slavonic languages 84, 179, 371–2, 373,
see also Russian
- Spanish:
- based creoles 411 n. 11, 412
 - diachronic change in 302, 399
 - synchronic variation in 25, 28 n. 4,
32, 54, 83, 326, 430
 - variants of 399, 430
- Swedish 45, 50, 54, 136, 138, 214, 220,
302 n. 6
- Swiss German 182, 340
- Tamil 411 n. 11
- Tigre 343–5
- Tigrinya 343–5
- Trinidad Creole 410
- Turkish 174
- Veneto (dialect of Italian) 38–9
- Welsh 45–8, 50, 54, 64, 75, 83, 400–4,
441 (fr)
- Welsh English 400–4, 441 (fr)
- West African languages 97,
420–1
- Yiddish 51 n. 15
- Yoruba 421, 422, 423