

# Against Synchronic Chain Shifting

Nicholas Charles Neasom

Department of Linguistics  
University College London

Thesis submitted in partial fulfilment of the requirements for  
the degree of Doctor of Philosophy

UCL

September 2016



## Declaration

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

*Nicholas Neasom*



## Abstract

In a synchronic chain shift, some underlying form /A/ is realized on the surface as a distinct surface form [B]. Simultaneously, underlying instances of /B/ are realized as a further distinct form [C]. This has traditionally been viewed as problematic; if /B/ cannot surface faithfully, then underlying /A/ should map to [C]. I argue that the  $A \rightarrow B \rightarrow C$  nature of synchronic chain shifts is not genuinely problematic. There is nothing substantive uniting the class of processes that exhibit an  $A \rightarrow B \rightarrow C$  mapping.

I begin by showing that whilst there are several theories that model synchronic chain shifting, there is no genuine consensus on how the term is to be defined. I compare chain shifts in synchrony to shifts in diachrony and acquisition, concluding that in both cases there are few genuine similarities. Next, I present a detailed survey of putative synchronic chain shifts. Building on a collection compiled by Elliott Moreton (2004a), I revise and update information on existing entries and add new examples.

Using examples from the survey as case studies, I argue that there are no coherent groups of shifts above the level of the segment. Furthermore, whilst there are coherent classes of processes at the level of the segment, the forces that motivate them are indifferent to whether they result in a chain shift. Finally, I present the results of a pilot Artificial Grammar Learning experiment. The experiment gives no reason to suppose that chain shifts are any more or less learnable than similar, non-chain shift patterns.

I conclude that there is nothing uniting the set of putative synchronic chain shifts. If we are to treat these effects as genuinely synchronic, effects featuring  $A \rightarrow B \rightarrow C$  mappings are better explained using more general principles of the phonological grammar.



## Acknowledgements

As I understand it, a PhD is essentially an accrual of debts. These debts may be financial, emotional, or spiritual in nature. I acknowledge some of my more prominent creditors here as a promissory note for, or in lieu of, payment.

Firstly, there is no way that I would have finished this thesis without John Harris. His invaluable guidance, wisdom, insight, and understanding have all been constants throughout these four years, and I am very grateful indeed for that. I am also grateful for all of the drinks.

Next, I would like to thank Andrew Nevins for introducing me to the concept of chain shifting during an inspiring class back in 2011, and for all of his help, advice, and generosity since then. I would like to thank Jamie White as well, for many useful and stimulating discussions, particularly about the experimental side of phonology.

I have been very lucky to be part of a PhD room filled with excellent people doing excellent work, and I will miss being a part of that. Some of these excellent people are Zoë Belk, Laura Aldridge, Florian Breit, Faith Chiu, Giulio Dulcinati, Elizabeth Eden, Patrick Elliott, Bruno Fernandes, Sam Green, Jiri Kaspar, Ya-Fang Lu, Johnny McIntosh, Emilia Molimpakis, Caterina Paolazzi, Lewis Pollock, Chao Sun, Irini Symeonidou, Gise Tomé, and Kevin Tang. I am sorry for being loud.

There are also some friends outside the PhD room to whom I am particularly beholden. I am thankful to Dr and Mrs Hine, Rob and Beth, and Sam and Simon, for living with me either for more than twenty years, or at exceptionally close quarters. Your forbearance is much appreciated. My parents have also been an unfailing source of support and encouragement throughout this process in particular and my life in general, and I state here that I am very lucky to have them.

I did not expect to spend part of my writing up year getting married twice, but that is what ended up happening and I am very glad it did. I want to thank Jill for her patience, for being a sounding board for some breathtakingly stupid ideas, and for all of the other wonderful things that she does and is.

Finally, I point out here that this research was funded by the Arts and Humanities Research Council, to whom I am very grateful.



# Table of contents

Declaration .....	3
Abstract .....	5
Acknowledgements .....	7
Table of contents .....	9
List of tables .....	15
List of figures .....	19
List of abbreviations .....	21
Chapter 1: Introduction .....	23
Chapter 2: Previous theoretical approaches to synchronic chain shifting.....	29
2.1 Some shared ground, and some ground rules.....	29
2.2 Rule-ordering .....	30
2.3 Scalar theories .....	32
2.3.1 <i>n</i> -ary rules.....	33
2.3.2 Ternary scales: Gnanadesikan (1997) .....	36
2.3.3 Abstract scales: Mortensen (2006).....	40
2.4 Local Conjunction .....	43
2.5 Contrast Preservation .....	45
2.6 Optimality Theory with Candidate Chains.....	52
2.7 Summary .....	55
Chapter 3: Chain shifts: Synchrony vs. diachrony and acquisition.....	57
3.1 Introduction .....	57
3.2 Diachronic chain shifts.....	57
3.2.1 How much should we be attempting to model? .....	58
3.2.2 The roots of synchronic chain shifting.....	62
3.3 Shifts in L1 acquisition .....	66
3.3.1 The reality of child chain shifts: Evidence from diary studies..	68
3.3.2 More general concerns .....	72
3.4 Summary .....	78
Chapter 4: The limits of chain shifting, and a new chain shift corpus.....	81

4.1 Introduction.....	81
4.2 The state of play.....	82
4.2.1 Moreton & Smolensky (2002) and Moreton (2004a).....	82
4.2.2 Other sources and entry criteria.....	84
4.2.3 The Revised Corpus of Putative Synchronic Chain Shifts: A User's Manual.....	86
4.3 The Revised Corpus of Putative Synchronic Chain Shifts.....	91
A-Hmao.....	91
Arabic (Bduul dialect).....	92
Arabic (Bedouin Hijazi dialect) 1.....	93
Arabic (Bedouin Hijazi dialect) 2.....	94
Arabic (Bedouin Hijazi dialect) 3.....	95
Basaa.....	96
Basque (Etxarri Navarrese dialect).....	97
Bengali (Standard colloquial dialect).....	98
Catalan.....	99
Chemehuevi.....	100
Chuckchi.....	101
Danish.....	102
Dutch (Hellendoorn dialect).....	103
English (African American Detroit dialect).....	104
English.....	105
English (Southern Standard British dialect).....	106
Finnish 1.....	107
Finnish 2.....	108
Gbanu.....	109
Hidatsa.....	110
Hmong (Dananshan dialect).....	111
Hmong (Shuijingping dialect) 1.....	112
Hmong (Shuijingping dialect) 2.....	113

Hmong (Xinzhai dialect).....	114
Icelandic .....	115
Inupiaq (Barrow dialect) .....	116
Irish 1.....	117
Irish 2.....	118
Italian (Servigliano dialect).....	119
Italian (Central and southern vernacular dialects).....	120
Japanese (Tokyo dialect).....	121
Karok.....	122
Kayardild.....	123
Kikuria.....	124
Manya.....	125
Misantla Totonac.....	126
Mohawk.....	127
Mwera.....	128
Nzɛbi .....	129
Nzema.....	130
Ojibwa .....	131
Palauan .....	132
Pero.....	133
Pipil .....	134
Polish.....	135
Polish (Kashubian dialect) .....	136
Sanskrit.....	137
Sea Dayak.....	138
Serbo-Croatian 1 .....	139
Serbo-Croatian 2 .....	140
Shoshone (Tumpisa Panamint dialect).....	141
Southeastern Pomo .....	142
Spanish (Gran Canaria dialect) .....	143

Spanish (Lena dialect) .....	144
Tarascan .....	145
Thok Reel .....	146
Toba Batak .....	147
Tzutujil (Santiago dialect) .....	148
Wintu 1 .....	149
Wintu 2 .....	150
Xiamen .....	151
Yagua .....	152
Yokuts (Wikchamni dialect) 1 .....	153
Yokuts (Wikchamni dialect) 2 .....	154
Yokuts (Yowlumne dialect) 1 .....	155
Yokuts (Yowlumne dialect) 2 .....	156
4.4 Trends in the corpus .....	157
4.4.1 Language family .....	157
4.4.2 Domain .....	159
4.4.3 Processes involved .....	160
4.4.4 Shifts as a whole .....	164
4.5 Where do we go from here? .....	166
Chapter 5: Reasons to doubt the existence of synchronic chain shift .....	169
5.1 Introduction .....	169
5.2 Practical problems .....	170
5.2.1 A lack of synchrony: Kashubian, African American Detroit English .....	170
5.2.1.1 Kashubian .....	170
5.2.1.2 African American Detroit English .....	173
5.2.2 Practical issues: Insufficient or unclear data .....	174
5.2.2.1 Basque (Etxarri Navarrese and otherwise) .....	174
5.2.2.2 Pero .....	178
5.2.3 Leaps into the unknown: Chain shifts and Richness of the base .....	179

5.3 Domain restriction.....	185
5.3.1 Two kinds of counterfeeding, ‘environments’ and ‘sources’..	185
5.3.2 A particularly problematic case: Sea Dayak .....	189
5.3.2.1 Problem 1: The data itself.....	190
5.3.2.2 Problem 2: Counterfeeding/chain shift analyses of the data are not insightful.....	191
5.3.2.3 Problem 3: The schematic.....	192
5.3.2.4 Two intrinsically ordered accounts of Sea Dayak....	195
5.3.3 No shared environment, no A → B → C mapping: Icelandic	201
5.3.4 When unification gets too unified: A seeming vowel deletion chain shift in Hidatsa.....	205
5.4 Chain shifting at the level of the segment: Chain shift as a smaller part of something bigger.....	212
5.4.1 Mutation and metaphony: rationale for treating them together	214
5.4.2 Chain shift as a smaller part of something bigger.....	218
5.4.2.1 Lenition mutation in Irish.....	219
5.4.2.2 Vowel lowering mutation in Thok Reel.....	222
5.4.2.3 Chain shift metaphony as an epiphenomenon of rule- based approaches.....	228
5.4.2.4 OT accounts of metaphony – Chain shifting has to be special, and that is a problem .....	236
5.4.2.5 Representations .....	241
5.4.2.6 Mutation and metaphony as morphology.....	246
5.5 Conclusion.....	250
Chapter 6: The learnability of synchronic chain shifts: Directions for future research and a pilot study .....	253
6.1 Introduction .....	253
6.2 Why vowel height?.....	257
6.3 Why artificial grammar? .....	259
6.4 The study .....	264
6.4.1 Stimuli .....	266
6.4.1.1 An AX test measuring consonant discriminability ..	268

6.4.1.2 Conditions and predictions .....	270
6.4.1.2.1 Simplicity vs. complexity .....	272
6.4.1.2.2 Coherence vs. incoherence: Does the pattern act on a natural class? .....	272
6.4.1.2.3 Transparency vs. ‘opacity’ .....	273
6.4.2 Participants .....	275
6.4.3 Methodology .....	275
6.4.3.1 Sub-studies .....	277
6.5 Results and discussion .....	279
6.5.1 Results for non-multiclickers .....	287
6.6 More general potential reasons for the null result .....	301
6.7 Further studies .....	307
6.8 Some speculations: What would it mean if the results were replicated? .....	310
Chapter 7: Conclusions .....	313
7.1 Introduction .....	313
7.2 Main findings .....	313
7.3 Implications .....	316
7.4 Concluding remarks .....	317
Appendix A: A reanalysis of Mascaró’s typology of Italian metaphony dialects .....	319
Appendix B: Stimuli for the AGL pilot study .....	323
References: .....	327

# List of tables

## Chapter 2

Table 2-1: Counterfeeding order ( $B \rightarrow C$ precedes $A \rightarrow B$ , chain shift results) .....	30
Table 2-2: Reverse order ( $A \rightarrow B$ precedes $B \rightarrow C$ ).....	31
Table 2-3: Gnanadesikan's analysis of Lena Spanish.....	38
Table 2-4: Mortensen's account of $a \rightarrow e \rightarrow i$ .....	41
Table 2-5: Martínez-Gil account of Lena Spanish raising .....	44
Table 2-6: Potential scenarios for the Lena Spanish shift.....	47
Table 2-7: Lena in Contrast Preservation.....	48
Table 2-8: Potential chain shifts from Gurevich's lenition corpus .....	52
Table 2-9: OT-CC account of Lena (1).....	53
Table 2-10: OT-CC account of Lena (2).....	54

## Chapter 3

Table 3-1: Fulcrand's analysis of the GEVS.....	60
Table 3-2: Ettlenger's illusory chain shift .....	69
Table 3-3: Relative rates of application of the variants for word-final /s/ and /ʃ/ in Z's stage 8 .....	70
Table 3-4: Tableaux of a Local Conjunction analysis of the horse-shoe shift.....	71

## Chapter 4

Table 4-1: Language families in the corpus .....	158
Table 4-2: A cleaned table of language families in the corpus .....	158
Table 4-3: Domains of chain shifting.....	159
Table 4-4: Raw count of processes in the table.....	161
Table 4-5: Processes from Moreton's corpus.....	163
Table 4-6: Unified processes in the corpus .....	166

## Chapter 5

Table 5-1: Kirchner's model of Etxarri Navarrese Basque .....	176
Table 5-2: Underlying /ai/ on Gnanadesikan's analysis .....	182
Table 5-3: Underlying /e/ on Gnanadesikan's analysis .....	182
Table 5-4: The Kenstowicz & Kisseberth analysis of Sea Dayak .....	189
Table 5-5: Slightly adapted tableaux for Sea Dayak .....	191
Table 5-6: Phonology: Nasal harmony applies wherever possible.....	197
Table 5-7: Stratum 1 .....	199
Table 5-8: Stratum 2 .....	199
Table 5-9: Local Conjunction tableaux for Hidatsa .....	207
Table 5-10: Radical and eclipsed forms of consonants in Irish.....	216
Table 5-11: List of lenition alternations .....	219
Table 5-12: Underlying vowels and their surface realizations in Thok Reel .....	223
Table 5-13: Derivation for the incorrect, counterfeeding order .....	235
Derivation for the correct, feeding order .....	235
Table 5-15: Chain shifting dialects under Mascaró's analysis .....	239
Parkinson's height feature table for Servigliano Italian .....	241
Table 5-17: Parkinson's analysis of Servigliano .....	242
Table 5-18: Trommer's sonority scale.....	243
Table 5-19: Trommer's lenition analyses .....	244

## Chapter 6

Table 6-1: Examples of AX test stimuli .....	268
Table 6-2: Front vowels.....	269
Table 6-3: Back vowels .....	269
Table 6-4: Chain Shift examples .....	271
Table 6-5: Coherent Raise examples .....	271
Table 6-6: Mixed Raise examples .....	271
Table 6-7: Comparison of conditions .....	272
Table 6-8: Derivations in the three conditions .....	274

Table 6-9: Number of stimuli presented in the training phase by condition.....	276
Table 6-10: Overall results (proportions of correct answers) .....	280
Table 6-11: Model featuring full results .....	284
Table 6-12: ANOVA comparing models with and without Condition as a predictor.....	284
Table 6-13: Model of dynamic results .....	287
Table 6-14: Average scores for non-multiclickers .....	288
Table 6-15: Differences in overall score between multiclickers and nonmulticlickers .....	289
Table 6-16: Overall statistical analysis of non-multiclicking participants.....	291
Table 6-17: Modelling the interaction between static and dynamic mappings amongst non-multiclickers.....	292
Table 6-18: Static mapping results for non-multiclickers.....	293
Table 6-19: Overall results from small control study by vowel.....	295
Table 6-20: Conservative participants by Condition.....	296
Table 6-21: Conservative and anti-conservative participants by condition .....	296
Table 6-22: Strategic participants from complete results.....	297
Table 6-23: Overall scores per vowel for non-multiclickers.....	297
Table 6-24: Sub-study 1 results.....	300
Table 6-25: Scores with ceiling and strategic participants removed.....	303
Table 6-26: A comparison with previous studies.....	310

## Appendix A

Table A-1: Mascaró's analysis of neutralization.....	319
Table A-2: Neutralization analysis without contrast preservation constraint .....	319
Table A-3: Mascaró's analysis of dialects with no low-mid metaphony .....	320
Table A-4: No low-mid metaphony analysis without contrast preservation constraint .....	320
Table A-5: Mascaró's analysis of a dialect that creates low-mid diphthongs.....	321
Table A-6: Low-mid diphthong dialect analysis without contrast preservation constraint (input /ɛ/) .....	321
Table A-7: Low-mid diphthong dialect analysis without contrast preservation constraint (input /e) .....	321

Table A-8: Mascaró’s analysis of a chain shifting dialect.....321

Appendix B

Table B-1: Stem vowel /a/ .....323

Table B-2: Stem vowel /e/ .....323

Table B-3: Stem vowel /i/ .....324

Table B-4: Stem vowel /o/ .....324

Table B-5: Stem vowel /u/ .....325

# List of figures

## Chapter 3

Figure 3-1: The approximate time course of the GEVS, according to Lass (1999)... 61

## Chapter 4

Figure 4-1: Screen-grab from Moreton's corpus..... 83

Figure 4-2: Processes involved in the corpus..... 162

Figure 4-3: Processes involved in Moreton's corpus ..... 164

## Chapter 5

Figure 5-1: Kashubian vowels..... 172

Figure 5-2: Standard Polish vowels ..... 172

Figure 5-3: The organization of the phonological grammar ..... 196

Figure 5-4: Possible lenition trajectories from Lass 1984..... 221

Figure 5-5: Acoustic measurements of Thok Reel vowels from Reid 2010 ..... 225

## Chapter 6

Figure 6-1: Xiamen tone circle..... 260

Figure 6-2: Experiment conditions..... 265

Figure 6-3: Proportion of correct answers by vowel pair..... 270

Figure 6-4: Overall proportions of correct results x mapping..... 280

Figure 6-5: Boxplots illustrating spread for static, dynamic and total results..... 282

Figure 6-6: Non-multiclick results ..... 288

Figure 6-7: Boxplots illustrating Static, Dynamic and Total scores for nonmulticlickers ..... 290

Figure 6-8: Element theoretic treatment of the vowel system ..... 304



## List of abbreviations

ACC = Accusative

AGL = Artificial Grammar Learning

DAT = Dative

EEG = Electroencephalography

GEVS = Great English Vowel Shift

GP = Government Phonology

IMPER = Imperative

IMPERF = Imperfective

INFIN = Infinitive

LC = Local Conjunction

NEG = Negation

NOM = Nominative

NZES = New Zealand English Shift

OT = Optimality Theory

OT-CC = Optimality Theory with Candidate Chains

PERF = Perfective

PL = Plural

RECIP = Reciprocal

SPE = *The Sound Pattern of English* (Chomsky & Halle 1968)

SG = Singular

RotB = Richness of the Base

UTIL = Utilitive



## Chapter 1: Introduction

In this thesis, I argue that a coherent, unified class of synchronic chain shifts does not exist, and that the phonological grammar is indifferent to whether or not its operation results in what have been described as chain shift patterns. In order to make this argument, it is important to be clear about the kinds of processes that are typically seen as synchronic chain shifts. An example comes from the Lena dialect of Spanish, spoken in the Asturias region. The masculine singular suffix {-u} conditions a set of related changes in tonic vowels. Instances of the low vowel /a/ raise to [e], as in *gata* ‘cat’, (fem.sg) vs. *getu* ‘cat’, (masc.sg). Simultaneously the mid vowels /e/ and /o/ raise to their high counterparts, [i] and [u], as in *nena* ‘girl’ vs. *ninu* ‘boy’ (see, e.g. Hualde (1989) for further description). However, an important question immediately suggests itself. Clearly, there is some proscription against underlying low vowels surfacing faithfully in this context. It is odd, however, that they should surface as mid vowels, given that *underlying* instances of mid vowels do not surface faithfully.

This raises two related questions, if we make the prevalent assumption that both of these vowel raising effects are active, synchronic processes (see, e.g., Gnanadesikan (1997), Moreton & Smolensky (2002), Martínez-Gil (2006)). Firstly, if mid vowels are licit surface forms in derived environments, why should underlying mid vowels be realized as high vowels in these contexts? Secondly, if underlying mid vowels are not realized faithfully, why are underlying low vowels not realized as high? In short, why does Lena Spanish have *ninu* as opposed to *\*nenu*, or *getu* instead of *\*gitu*? Accounts of synchronic chain shifting have focused more on this second question (see Gnanadesikan (1997), Moreton & Smolensky (2002), and so on, but also see Lubowicz (2003a) for a more holistic account). However, the general idea is almost certainly that solving one also solves the other.

In a recent review article, Anna Lubowicz gives the following definition of chain shift: “underlying /A/ maps onto surface [B] and underlying /B/ maps onto surface [C] in the same context, but crucially, underlying /A/ does not become [C]. Thus a chain shift has a standard representation  $A \rightarrow B \rightarrow C$ ” (2011, p. 1717). This

definition is meant as a catchall that is supposed to be as effective for chain shifts in diachrony, language acquisition, or dialectal variation as it is for shifts in synchronic grammar. In chapter 3, I discuss the implications of this lumping together, which I argue are extensive. For now, however, the definition will suffice, particularly as it accords well with other contemporary definitions of the term (see those in, e.g. Moreton & Smolensky (2002), Mortensen (2006), Mahanta (2012)). Each of these definitions is slightly different, influenced by the theoretical positions I discuss in subsequent sections, but they are all similar at their core.

There are many processes in the world's languages that fit the superficial  $A \rightarrow B \rightarrow C$  mapping suggested by Łubowicz. I argue that this, in and of itself, is not of any great linguistic significance. Close examination of individual  $A \rightarrow B \rightarrow C$  mappings fails to reveal a common motivation for the set of putative chain shifts. This collection of processes is smaller than has been previously reported, as there are many cases where the data does not justify the postulation of an  $A \rightarrow B \rightarrow C$  mapping. Even among the set of processes where it is fair to postulate such a mapping, the set of processes is too diverse to be modeled by a 'one-size-fits-all' theory. Moreover, there is no reason to suspect that the lack of an  $A \rightarrow C$  mapping, which is seen by most chain shift theorists as the key problem in discussions of chain shifting, is ever anything other an epiphenomenon caused by the regular operation of phonology and/or morphology.

Indeed, noting that many phonological processes result in an  $A \rightarrow B \rightarrow C$  mapping and suggesting that this may relate the processes in some way has unfortunate consequences. It suggests that these effects represent a unified class of processes, requiring a specific, productive, phonological solution that will work in all cases. As we will see in the following chapter, many of these solutions involve adding to the theoretical machinery of phonological grammar, expanding its power in this way or that, in order to cope with this somewhat nebulous, apparently recalcitrant set of data. I argue that whilst many of the theories that have been derived to account for chain shifting are ingenious, they are more reflections of the complex knowledge that linguists have of language than they are realistic representations of the tacit knowledge of speakers. By failing to attend to the asymmetry between what linguists know about language and the actual knowledge that speaker/hearers have, we may

end up explaining things that we do not need to explain, or things that we do not yet really understand. Whenever the object requiring explanation is something abstract, problems of over-explanation and conflicts over how to define important concepts can arise. That there are multiple definitions in the literature of what constitutes a synchronic chain shift only makes direct comparison that much more difficult.

This thesis discusses the meta-theoretical issues laid out above, in relation to the specific empirical problem of how to characterize and represent effects that have been called synchronic chain shifts. In the following chapter, I discuss previous and current approaches to these problems, before beginning a more thoroughgoing investigation into processes that have been claimed to be synchronic chain shifts.

The evidence presented in the following chapters strongly suggests that there is no unified, coherent set of synchronic chain shifts. However, in general, synchronic shifts are assumed to exist and be problematic in most rule-based and Optimality-Theoretic approaches to phonology. The aim of this thesis is to question this assumption. To this end, it is necessary to break the question of synchronic chain shifting down into the three sub-questions in (2).

- (2) (a) Can synchronic chain shifts be disentangled from shifts in diachrony and acquisition?
- (b) Is there a canonical set of synchronic chain shifts?
- (c) Are synchronic chain shifts learnable?

Each of these topics are addressed at length across the thesis. (2a) is the subject of chapter three, (2b) is discussed in chapters four and five, and (2c) is addressed in chapter six. I give brief introductions to each of these chapters below.

### *Chapter 3: Can synchronic chain shifts be disentangled from shifts in other domains?*

The majority of work on chain shifting comes from a diachronic perspective (see, e.g., Martinet (1955) and Labov (1994)). There are studies that use the methods of synchronic analysis, and apply them to diachronic effects (e.g., Miglio & Morén (2003), Noske (2012), Fulcrand (2015)), but whether there is any substantive

relationship between chain shifts in synchrony and diachrony is still an open question.

In the synchronic literature, one of two positions is generally adopted. The first is exemplified by Kirchner, who states that the analysis that he uses to model a vowel raising shift in Nzèbi would be unnecessary in a diachronic analysis, “since the ordering of sound changes presumably corresponds to distinct historical stages” (1996, p. 341). Łubowicz (2011) takes the reverse position. Her definition of chain shifting includes shifts in diachrony and synchrony (and also in acquisition, and dialectal variation).

There is a similar division in the literature on first language acquisition. On the one hand there are scholars who assume that child chain shifts are problems that can be equated to chain shifts in the synchronic adult grammar, and can be addressed with similar representational machinery (e.g., Dinnsen & Barlow (1998), Jesney (2005; 2007), Dinnsen et al. (2011)). On the other, there are theorists who dispute that child chain shifts are genuinely problems of grammatical competence (e.g., Hale & Reiss (2008), Smith (2010), Walton (2012)). Under these approaches it would not be possible to genuinely equate shifts in acquisition with shifts in the stable adult grammar. I argue that the similarities that occur across chain shifting in synchrony, diachrony, and acquisition are superficial at best, and non-existent at worst. There is, in my estimation, no substantive connection between shifts across these three domains.

#### *Chapters 4 and 5: Is there a canonical set of synchronic chain shifts?*

The findings of chapter 3 enable a narrowing of focus, in subsequent chapters, to synchronic putative shifts. Chapter 4 presents a large, detailed database of such effects, called the Revised Corpus of Putative Synchronic Chain Shifts, based mainly on a reanalysis of a compendium of chain shifts compiled by Elliott Moreton (2004a). I discuss the methodology of sampling, present the corpus, and finish by analyzing salient trends. This chapter essentially acts as a resource for anyone interested in the study of synchronic chain shifting, and is to my knowledge the largest and most detailed collection yet available of these effects. For the reader who is interested primarily in the overall arguments of the thesis, the most relevant

section of this chapter is section 4.4, which discusses the trends in the corpus. The rest of the chapter can be skipped without undue damage to the narrative of the thesis.

In chapter 5, I first present case studies of putative shifts that can be argued to be misdescribed, based on scant or flawed data, or artefacts of controversial theoretical assumptions. I then winnow down the corpus in two further ways. Firstly, I argue for a domain restriction to the level of the segment (in line with, for example, McCarthy (2007)) by showing that whilst there are coherent classes of apparent chain shift that apply time and again at the level of the segment, this is not true above that level. Even when there are apparent trends above the level of the segment, I argue that the processes involved do not constitute plausible candidates for shifthood. In many of these cases, there may not even be a genuine  $A \rightarrow B \rightarrow C$  order when the overall effect is considered. Even when there is such a mapping relation, it is often the only logical consequence of the regular operation of either genuinely intrinsic phonological ordering, or the logical result of morphology applying once and once only.

I then examine cases where chain shifts occur as parts of larger effects. I argue that in such cases, no theory of chain shift that rests on blocking an  $A \rightarrow C$  mapping can give a plausible account of the entire effect. Furthermore, I argue that there is a simpler solution in these cases; that they constitute unified effects that are indifferent to whether chain shifts are created or not. I discuss how this undercuts the notion of a chain shift as a problem that needs solving, and question exactly what it is about chain shifts that is especially troubling for theoretical phonology. I consider approaches in which chain shifts are less problematic, and discuss the possibility that certain chain shift effects, particularly mutation and metaphony effects, may not even be phonological processes at all. Chapter 5 concludes by discussing the idea that it is difficult to assess whether the phonology is actually involved in certain chain shift effects if one takes a purely theoretical perspective.

*Chapter 6: Are synchronic chain shifts learnable?*

The final chapter of this thesis illustrates how one might investigate the empirical question of whether chain shifts are learnable. If chain shifts are to be considered genuine, productive processes in synchronic grammar, we would expect them to be harder to learn than similar processes that are less phonologically complex, and easier to learn than similar processes that are more phonologically complex. There are already experiments on the productivity of individual chain shifts, which typically take the form of wug-tests. For example, Zhang, Lai & Sailor (2006) study a circular tone-shift in Xiamen, and Nagle (2013) tests the productivity of a vowel height shift, not dissimilar to that in Lena Spanish, in Bengali. The results of these experiments suggest that neither pattern is fully synchronically productive.

I show how one might take a wider approach to the question, presenting results from a pilot experiment using the Artificial Grammar Learning paradigm. This experiment investigates whether one of the most common chain shift patterns, two-step vowel raising, is learnable in an experimental environment, compared to minimally different vowel shift patterns that do not feature a chain shift. Whilst the experiment has certain methodological difficulties that mean the results are not conclusive, the overall finding appears to be that the three patterns are all roughly equally learnable.

I conclude by reasserting that I do not believe that there is any coherent, unified set of synchronic chain shifts. In the best case, all that such processes share is a superficial  $A \rightarrow B \rightarrow C$  mapping, and even this mapping relationship cannot be observed as frequently as has been suggested in the literature.

## Chapter 2: Previous theoretical approaches to synchronic chain shifting

### 2.1 Some shared ground, and some ground rules

As mentioned in the previous chapter, everyone who has attempted to deal with chain shifts has grappled with essentially the same problem. If there is some process in which /B/ surfaces as [C], and some reason why /A/ cannot surface faithfully, then it is necessary to rule out the /A/ → [C] mapping. This casts synchronic chain shifts as processes that are inherently problematic, and assumes that some kind of theoretical machinery is required to prevent the neutralization that would otherwise occur. In what follows I present the approaches taken to synchronic chain shift in several frameworks, beginning with rule-ordering approaches in serial phonology<sup>1</sup>. I then discuss various methods of modeling chain shifts that have been proposed in Optimality Theory (OT).

The overriding aim of this section is not simply to describe each approach to chain shifting, but also to discuss the effect that employing each method has on the phonological grammar as a whole. Particularly with regard to specific OT methods of modeling chain shift, there are wide-ranging consequences for the grammar that are often far from mainstream phonological views. This is important because all of the theories below can be shown to model chain shifts reasonably effectively. It is therefore the wider impact that they have on the phonological system that is of most importance. Throughout this section I will refer repeatedly to the Lena Spanish chain shift that was used as an example at the start of chapter 1. I show that all of the chain shift theories below can model the Lena Spanish shift. I choose this shift because it is well-studied (see, e.g. Hualde (1989), Gnanadesikan (1997), Martínez-Gil (2006)) and is considered to be a fairly ‘canonical’ kind of chain shift, as it is used as an example of chain shifting in these and other previous approaches (e.g. Parkinson (1996)). I give a full set of examples from the Lena Spanish shift in (1) (examples from Gnanadesikan (1997, p. 209), citing Hualde (1989)).

---

<sup>1</sup> Certain chain shifts, particularly metaphonic vowel raising chains in Romance languages, have been analysed in great detail using non-linear rules (e.g., Kaze (1989), Cole (1998), Puente (1996), Calabrese (2011)). These approaches do not have the generality of the other approaches that are discussed in this chapter. I do discuss approaches like this in section 5.4.2.3.

(1)	Fem sg.	Masc sg.	Gloss
	<i>a</i> → <i>e</i>		
	gáta	gétu	‘cat’
	sánta	séntu	‘saint’
	ǰánu	ǰénu	‘diligent worker’
	blánka	blénku	‘white’
	páǰara	péǰara	‘bird’
	<i>e</i> → <i>i</i>		
	néna	nínu	‘child’
	kordéro	kordíru	‘lamb’
	séka	síku	‘dry’
	<i>o</i> → <i>u</i>		
	tsóba	tsúbu	‘wolf’
	kóǰa	kúǰu	‘cripple’
	silikótika	silikútiku	‘suffering from silicosis’

## 2.2 Rule-ordering

Perhaps the simplest way of modeling a synchronic chain shift is to use serially-ordered rules in a counterfeeding order, as pointed out by (at least) Moreton & Smolensky (2002) and Łubowicz (2011). The two key components to a rule-based analysis are the breaking down of the chain shift effect into two parts (an  $A \rightarrow B$  rule and a  $B \rightarrow C$  rule), and the extrinsic ordering of these rules so that the  $B \rightarrow C$  rule *precedes* the  $A \rightarrow B$  rule. A simple example of how this works, using Lena Spanish forms from above, is shown in (2) and table 2-1.

- (2) (a)  $A \rightarrow B$  rule: [+low] → [-low] / \_\_{-u}  
 (b)  $B \rightarrow C$  rule: [-low, -high] → [+high] / \_\_{-u}

Table 2-1: Counterfeeding order ( $B \rightarrow C$  precedes  $A \rightarrow B$ , chain shift results)

UR	/gat + a/	/gat + u/	/nen + a/	/nen + u/
[-low, -high] → [+high] / __{-u}	-	-	-	nínu
[+low] → [-low] / __{u}	-	getu	-	-
SR	[gata]	[getu]	[nena]	[ninu]

The crucial datum in table 2-1 is /gat + u/. Because the B → C rule applies first, /gat + u/ does not meet the structural description for a process that would undergo it, as the /a/ vowel in /gat + u/ is still specified as [+low]. Therefore, it only undergoes the A → B rule, which takes place later in the derivation. The consequences for reversing this order are shown in table 2-2.

Table 2-2: Reverse order (A → B precedes B → C)

UR	/gat + a/	/gat + u/	/nen + a/	/nen + u/
[+low] → [-low] / __{-u}	-	getu	-	-
[-low, -high] → [+high] / __{-u}	-	gitu	-	ninu
SR	[gata]	*[gitu]	[nena]	[ninu]

Again, consider /gat + u/. If the rule A → B is ordered *before* the B → C rule then both rules apply. /gat + u/ is first transformed into [getu], as a result of the A → B rule. Then, the B → C rule applies and [getu] becomes the unattested surface form \*[gitu]. This is, to be sure, a simple and elegant way of representing chain shifts, and gives a simple reason for the lack of an A → C mapping. There is no rule that takes underlying /A/ all the way to surface [C], and the B → C rule occurs before /A/ forms are changed to [B]. Therefore, the mapping is blocked.

There are, however, criticisms about the nature of the rules that make up the chain shift and the extrinsic ordering that is required to make the rules work correctly. The argument against the rules themselves has been made by, amongst others, Łubowicz (2003a; 2011; 2012). It is based on the premise that chain shifting is, at least sometimes, a unified process. If this is true, runs the argument, it is perhaps unwise to use two rules to express two reflexes of what is one process. Intuitively, this makes sense. Consider the derivation in table 2-1. The raising of low to mid vowels, and the raising of mid to high vowels happens in exactly the same place, for exactly the same reason. Łubowicz herself argues that not all shifts are genuinely unified, and that some shifts are two independent processes that happen to interact (she calls these ‘regular’ shifts). The problem with a counterfeeding order for *these* kinds of effects is that there is no a priori reason for the extrinsic ordering that would yield a

chain shift. As Smith (2010, p.2) notes, citing Miller & Chomsky (1963), the concept of extrinsic rule ordering poses a considerable learnability challenge to a speaker/hearer acquiring a language.

The rabbit hole of this problem probably actually goes deeper than the criticism that Łubowicz makes, if one considers traditional assumptions about how particular rule-orderings come about. Whilst acknowledging that there are exceptions, Kiparsky (1970/1982) proposes that each new rule that is added to the phonological grammar is ordered *after* every other rule that is already part of the phonology (see also Kiparsky (1973a), Bromberger & Halle (1989)). This is the main source of opaque rule orders, such as the counterfeeding order that would be necessary to model the Lena Spanish shift. According to Kiparsky, there is very little *a priori* justification for a such an order to occur otherwise, as normally “[r]ules tend to be ordered so as to become maximally transparent” (1971/1982, p. 75). This suggests that if both the  $A \rightarrow B$  and  $B \rightarrow C$  rules came into the grammar at the same time, there would be no justification for a counterfeeding order. However, this makes the prediction that chain shifts are never temporally unified processes, because the  $B \rightarrow C$  rule would necessarily have come into the grammar generations before the  $A \rightarrow B$  rule (see 5.4.2.3.1 for further discussion of these issues).

In sum, counterfeeding rule orders are capable of representing synchronic chain shifts, but they make troubling predictions about how such processes come into the grammar. The simplest explanation for shifts in such accounts is that the counterfeeding order is the result of chronology. However, this suggests that chain shifts in synchrony are not necessarily unified processes. On the other hand, if both parts of the process are simultaneous additions to the language, it is unclear how or why the phonology would split the process in two.

### **2.3 Scalar theories**

Depending on one’s perspective, a key advantage or major drawback with explaining chain shifts via counterfeeding rule orders is that there is no real limit on the kind of processes that can be involved. There are some chain shift theorists who have postulated that the only true shifts, like the Lena Spanish shift, are those in which both the  $A \rightarrow B$  and  $B \rightarrow C$  rules have the same motivation. This shared motivation

is explicitly encoded in the grammar, by assigning each structure involved in the shift a value in a scale of some sort. Below, I discuss three scalar theories. The first is based on rules and used by a wide variety of phonologists. This is followed by discussion of two constraint based theories, one posited by Gnanadesikan (1997), and one by Mortensen (2006).

### 2.3.1 *n*-ary rules

An alternative to rule-ordering that does not encounter the problems discussed in the previous subsection is some kind of rule that is multi-valued, or *n*-ary. There are many proposals for modeling phonological processes that include multi-valued rules, for example those in Contreras (1969), Ladefoged (1971), Williamson (1977), and Lindau (1978). More specifically, *n*-ary rules have been proposed as a way of modeling chain shifts in progress (by Labov (1994, p.225), based on Labov, Yaeger, and Steiner (1972)) and synchrony (by, for example, Frajzyngier (1989) and Andersen (1993))<sup>2</sup>. An example of such a rule that would model the Lena Spanish shift would essentially take the form shown below (the specific *n*-ary conventions are borrowed and adapted from Frajzyngier (1989, p. 40)):

$$(3) \quad V \rightarrow [+1 \text{ Height}] / \_ \{-u\}$$

At a glance, this looks like it solves the problem of modeling a unified process with two rules. However, the value of *V* is relative. Even though the operation of adding a degree of height can be argued to be unchanging and invariant, the result of this operation will be different in a substantive way depending on the input. As an illustration, consider a simple final devoicing rule, where any voiced obstruent loses its voicing word-finally:

$$(4) \quad [-\text{son}, +\text{voi}] \rightarrow [-\text{voi}] / \_ ]_{\text{wd}}$$

---

<sup>2</sup> It should be pointed out that neither Frajzyngier nor Andersen refers to the processes that they model as chain shifts in their own analysis. However, Moreton (2004a) lists the effect discussed by Frajzyngier as a chain shift, and Trommer (2011) labels the effect discussed by Andersen as a chain shift.

A language with the rule in (4) and the voiced obstruent inventory /b d g z/ would have the surface forms [p t k s] word-finally once the rule had operated. But it can be seen that (4) performs exactly the same operation on /b/, /d/, /g/, and /z/, changing only the feature that they share, [+voice].

Considering again our potential *n*-ary rule in (3) and the Lena Spanish vowel system /a e i o u/, it is first necessary to carve the vowel system up into three distinct degrees of height. This would then mean that /a/ would have a height of 0, whilst /e, o/ would have a height value of 1 and /i, u/ would have a height value of 2. An immediate problem with a rule of this nature, at least with regard to Lena, is that it must be stipulatively constrained not to apply when the height value of the underlying form is 2, as there are no vowels in Lena that are higher than /i, u/.

The stipulative nature of this analysis actually goes somewhat further. In Frajzyngier's discussion of Pero, despite the existence of many phonological processes that act on vowel height, only one process uses *n*-ary features as opposed to the binary features used for all other processes adumbrated in Frajzyngier's grammar, whether or not vowel height is involved. This implies that the use of an *n*-ary rule is an ad hoc reaction to a recalcitrant process, rather than a principled advocacy of the scalar way of thinking (see section 5.2.2.2 for further discussion of Pero).

Another, deeper problem concerns the (at least) ternary nature of such a system. Whilst there are approaches to phonology that embrace the idea of ternary computation in a meaningful way (see the next subsection), most mainstream theories of phonology still apply a binary limit to computation. For example, in two well-known undergraduate textbooks, binarity of features is assumed (Hayes (2009), Gussenhoven & Jacobs (2011, p. 74)). The authors go on to state that there are now also privative, or unary, accounts of distinctive features. However, to my mind this is still philosophically closer to the standard conception of features than ternarity, in that what is at issue is still presence vs. absence, rather than where a particular segment is positioned on some scale. Furthermore, I argue that allowing exceptions to binarity does not just do philosophical damage. If height is expressed through *n*-ary features, but all other phonetic dimensions, for example backness, are expressed

in binary features, this creates a fundamental mismatch in the system. A system which is binary is not simply an  $n$ -ary system with only two levels, at least in its original conception. Jakobson & Halle (1956) discuss binary features as a system of *oppositions*:

“Each of the distinctive features involves a choice between two terms of an opposition...In a message conveyed to the speaker, every feature confronts him with a yes/no decision” (p.5)

This means that a system proposed by, for example, Contreras (1969), in which “pluses and minuses may be conceived of as integers for scales having two terms” (p.4), changes the grammar entirely. The vast majority of the scales in such a grammar would only have two points, or would have to be replaced by completely different features. In fact, there are many processes involving vowel height which are non-scalar, thus actively harder to model using  $n$ -ary features. For example, consider a vowel lowering process in Buchan Scots discussed by Paster (2004). In certain contexts, the high vowel /i/ lowers to [e] when it follows any non-high vowel (examples of surface forms from Paster (2004, p.365)):

- (5) (a) vere ‘very’
- (b) mēne ‘many’
- (c) lɜle ‘lily’
- (d) mʌne ‘money’
- (e) kafe ‘coffee’
- (f) glɔre ‘glory’
- (g) sɔre ‘sorry’

Paster characterizes the process through the rule in (6) (p.367):

- (6) [v, -back] → [-high] / [V, -high] X<sup>3</sup> \_\_

---

<sup>3</sup> X stands for a fairly complex set of Buchan Scots consonants.

Using binary features, the vowel lowering process is easily captured. Unstressed front vowels lose their [+high] specification, turning /i/ into [e]. It is possible to capture the wide set of triggering vowels (/e ε ɜ ʌ a o ɔ/) with the simple specification [-high]. If height is scalar, by contrast, there is no unifying height feature that binds together, for example, low and mid vowels. In a system like Buchan Scots, which in Paster’s analysis has four distinct heights<sup>4</sup>, there are three separate kinds of vowel that trigger the lowering process; [0 height], [1 height], and [2 height]. The only way to simplify this expression would be a rule like that in (7).

$$(7) \quad [3 \text{ height}] \rightarrow [2 \text{ height}] / [\neq 3 \text{ height}] \text{ X } \_\_$$

Clearly we do not want this kind of Frankenstein rule, where an *n*-ary feature can also be specified negatively, in the phonological grammar.

An advantage of using binary rules to carve up a multi-valued phonetic space is that it is an easy matter to define multiple natural classes within that space. This idea has been present in phonology since at least W. Wang’s (1967) work on tone features, where he explicitly rejects *n*-ary rules, despite evidence suggesting the existence of five separate height values for tone in “some Black Miao languages” (p. 96). In our example above, the feature [-high] gives us a clear natural class. Assigning every vowel height its own unique height feature means that, without ad hoc conditions on rules or environments, there is no way to get across the notion that vowels of different heights may pattern together in phonological processes.

### 2.3.2 Ternary scales: Gnanadesikan (1997)

Gnanadesikan (1997) takes a significantly different view of phonology not only to traditional binary feature approaches, but also to the unconstrained *n*-ary scales just discussed. Gnanadesikan is primarily interested in the behaviour of voiced obstruents, fricatives, and mid vowels. She notes that, crosslinguistically, the behaviour of these elements suggests that they are the midpoints of ternary scales (terminology from Gnanadesikan (1997, pp. 1-2)).

---

<sup>4</sup> Height classes in Buchan Scots, according to Paster (p.361) are /i, u/, /e, o/, /ε, ɜ, ʌ, ɔ/ and /a/

- (8) (a) Inherent voicing scale:  
 Voiceless obstruent >> Voiced obstruent >> Sonorant
- (b) Consonantal stricture scale:  
 Stop >> Fricative, liquid >> Vowel, laryngeal
- (c) Vowel height scale:  
 HIGH >> MID >> LOW

Gnanadesikan's main reasoning for the organization of these properties on three-way scales is the question of why, in her words, "mid vowels sometimes act with high vowels, sometimes with low vowels, sometimes alone, and sometimes with both in a chain shift" (p.2). Gnanadesikan's scales make the predictions that any combination of adjacent elements on the scale can act together, and that non-adjacent elements may not act together. In her review of the scalar literature (which shares sources and content with the review above, though it is, understandably, less critical of the scalar enterprise) Gnanadesikan gives particular credit to the work of Foley (1970) for informing her scalar model.

Foley's conception of phonology is somewhat similar to modern 'substance-free' phonologists (e.g., Hale & Reiss (2008), Blaho (2008)), typified by his assertion that "Physical data is part of a system of physics. Phonological data is part of a system of phonology" (p.89). With this in mind, the goal of phonology is to describe the relations between segments, without any necessary commitment to phonetic fidelity. For example, Foley discusses lenition and fortition trajectories in a variety of languages, but with the assumption that strength is a relative property, only meaningful in terms of the relationships it explains, not by itself as a substantive concept (p.89). Gnanadesikan takes the notions of scalarity and ternarity from Foley (and others), but with a greater commitment to phonetic substance, as is shown by the scales in (8a-8c). She also makes the salient point that whilst Foley's theory can model difficult processes like chain shifts, it would have trouble with the vast majority of non-scalar processes (p.9).

Gnanadesikan provides an analysis of the Lena Spanish shift using her scalar OT constraints, which I copy in table 2-3 (1997, p. 211). For completeness I have added

the form /saku/ to the second tableau, which is not included in Gnanadesikan’s original).

Table 2-3: Gnanadesikan’s analysis of Lena Spanish

gatu	IDENT-ADJ	IDENTICAL[HIGH]	ADJACENT[HIGH]	IDENT[VH]
gatu		*	*!	
☞ getu		*		*
gitu	*!			

seku	IDENT-ADJ	IDENTICAL[HIGH]	ADJACENT[HIGH]	IDENT[VH]
saku		*!	*	*
seku		*!		
☞ siku				*

The shift is modeled using four constraints, all of which work on the basic concept of similarity (essentially faithfulness, in OT terms). The candidate /gatu/ violates two of these constraints, IDENTICAL[HIGH], and ADJACENT[HIGH]. IDENTICAL[HIGH] is violated by forms containing any vowel that is not specified as HIGH.

ADJACENT[HIGH] is violated by any vowel that is more than one step on the scale away from HIGH. This means that whilst IDENTICAL[HIGH] is violated by both LOW and MID vowels, ADJACENT[HIGH] is more permissive. It is satisfied by MID vowels and only violated by LOW vowels. It is the violation of this constraint that rules out a faithful realization of underlying /gatu/.

Given that /getu/, the eventual winning candidate, also violates IDENTICAL[HIGH] (as well as the low-ranked general faithfulness constraint IDENT[VH]), there needs to be some mechanism to rule out the candidate /gitu/, which satisfies IDENTICAL[HIGH]. Gnanadesikan’s solution to the  $A \rightarrow C$  problem is a constraint that combines the principles of adjacency and general faithfulness, IDENT-ADJ. IDENT-ADJ is violated by any candidate that deviates from the input by more than one step on the scale. /gitu/ violates this constraint because on a three-point scale of vowel height,  $a \rightarrow e \rightarrow i$ , /a/ and /i/ are not adjacent. Given that IDENTICAL[HIGH] is highly ranked, and both [saku] and [seku] violate this constraint, the only possible realization is [siku].

Despite not going as far as Foley's proposal, Gnanadesikan's analysis still requires a complete reconceptualization of the phonological grammar, at the level of both underlying forms, and the way that constraints are able to function. OT is a system of computation and not representation, which means that, in a sense, issues such as whether features are unary, binary, ternary, or *n*-ary are unimportant. The computational system can handle all such feature systems. Indeed, Mortensen points out that the key change in the grammar induced by adopting Gnanadesikan's proposal is representational (2006, p. 6)). However, Gnanadesikan's proposal forces the computational system to function in a manner contrary to "the well-established generalization that linguistic rules do not count beyond two" (Kenstowicz, 1994, p. 597). Obviously, a ternary system requires a phonology that can count to three. Also, the constraints that Gnanadesikan uses require that phonological computation, as well as representation, be sensitive to all three points of the scale.

It is true that Gnanadesikan uses the language of traditional, non-ternary oppositions with her constraints, speaking of IDENTITY and ADJACENCY as guiding concepts. Additionally, Gnanadesikan's height features (HIGH, MID, and LOW) seem similar to unary features. However, to properly understand Gnanadesikan's metric of adjacency, it is necessary that the grammar is able to assign additional numeric values to these unary features, and perform operations of computation that understand that whilst 1 and 2 are adjacent, for instance, 1 and 3 are not.

Whilst Gnanadesikan shows that her scales are not just useful for explicitly scalar processes through a range of thorough case studies, the problem remains that there are many processes that involve phonological features such as nasality that are essentially impossible to make ternary. As Gnanadesikan acknowledges, consonant place is problematic for her theory. Whilst it is true that there are three major places of articulation, [labial], [coronal], and [dorsal], there is little to no evidence that there are any processes based on place that would be improved by suggesting that they form part of a scale. As Gnanadesikan says, "[t]here is no scalar behaviour in the move from one place to another" (p.228)<sup>5</sup>.

---

<sup>5</sup> Though see Ladefoged (1971) in support of [*n* place]

The consequence of this is a divided grammar, in which some oppositions are binary, and some are ternary. At first blush this may simply appear to be the same as models of distinctive features where some features are binary and others are unary.

However, unary features do not change the scope of the grammar in the same way that ternary features do; they do not necessitate that counting be added to the set of requirements needed for phonology to work.

### 2.3.3 Abstract scales: Mortensen (2006)

A more recent, and more radical, scalar theory is discussed in Mortensen (2006). Mortensen is primarily interested in tone sandhi effects in Min and Hmongic languages, and many of these have the  $A \rightarrow B \rightarrow C$  character familiar to chain shifts. However, Mortensen observes that in many cases, there does not appear to be anything phonetically natural about the relationship between A, B, and C. For example, in Lena Spanish, /a/ is low, /e/ is mid, and /i/ is high. There is a simple, unidirectional relationship between these three forms in terms of vowel height. However, consider the data in (9), from the Shuijingping dialect of Hmong (examples from Mortensen (2006, p. 85)).

- (9)  $HM \rightarrow \uparrow H$
- (a)  $ka^M$  ‘medicine’ +  $t\epsilon\epsilon^{HM}$  ‘liquor’  $\rightarrow ka^M t\epsilon\epsilon^{\uparrow H}$  ‘yeast’
- (b)  $na^L$  ‘bag’ +  $nt\epsilon^{HM}$  ‘book’  $\rightarrow na^L nt\epsilon^{\uparrow H}$  ‘book bag’
- $\uparrow H \rightarrow H$
- (c)  $he^M$  ‘chicken’ +  $he^{\uparrow H}$  ‘crow’  $\rightarrow he^M he^H$  ‘cock’s crow’
- (d)  $ti^L$  ‘RECIP’ +  $he^{\uparrow H}$  ‘curse’  $\rightarrow ti^L he^H$  ‘quarrel’

In the  $A \rightarrow B$  part of the shift, High-Mid tones (HM) rise to become superhigh ( $\uparrow H$ ) when following a mid (M) or low (L) tone in a compound. In the same context, underlying superhigh tones fall to become High (H), forming the  $B \rightarrow C$  part of the shift. It is apparent that in this situation, we do not have the clear trajectory that is present in shifts like Lena Spanish. The  $A \rightarrow B$  part of the shift is raising by two steps, and the  $B \rightarrow C$  part is lowering by one step. Mortensen states that in an earlier stage of the language, the Shuijingping shift was phonetically natural. Diachronic

changes have altered the character of the shift, but Mortensen still provides a synchronic analysis for this process, and others like it.

Like Gnanadesikan, in Mortensen’s conception of phonology, related elements are arranged on scales, though in Mortensen’s analysis there is no explicit ternary limit. Another key difference between Mortensen and Gnanadesikan’s accounts is that there is no requirement for the scales to conform to a particular order. For instance, Gnanadesikan’s vowel height scale must be ordered HIGH >> MID >> LOW. There is no provision for a reordering of this scale such as MID >> HIGH >> LOW, a commitment on Gnanadesikan’s part to a substantive link between phonology and phonetics. Mortensen acknowledges no such link.

To allow this amount of freedom in terms of how the scales can be formed, Mortensen’s constraints refer only to the abstract scale as a whole, and not any properties of the constituent parts of the scale. Mortensen exemplifies this with an abstract example of an  $a \rightarrow e \rightarrow i$  scale, which conveniently is exactly the same as the mapping of the Lena Spanish shift:

(10) Mortensen’s vowel height scale (2006, p. 70):

$$H = \{i\}_2 > \{e\}_1 > \{a\}_0$$

Table 2-4: Mortensen’s account of  $a \rightarrow e \rightarrow i$  (p.71)<sup>6</sup>

/a/	HIGHER(H)	SAME(H)	ENDMOST(H)	DIFF(H)
a	*!			*
e		*	*	
i		*	**!	

/e/	HIGHER(H)	SAME(H)	ENDMOST(H)	DIFF(H)
a	*!	*		
e	*!		*	
i		*	**	

The four constraints that are used in the tableau above all forbid certain scalar relations between inputs and outputs. With regard to the input /a/, the fully faithful

<sup>6</sup> I have edited the second of these tableaux slightly, as in Mortensen’s version, the output [e] for input /e/ violates SAME.

candidate is ruled out due to a violation of HIGHER, a constraint that forbids output realizations that are *not* higher on the scale than the input. This constraint essentially forces movement along the scale. The output [a] would also violate the lower-ranked constraint DIFFERENT, a constraint assigning a violation to any candidate that is on the same level of the scale as the input.

The candidates [e] and [i] both violate the constraint SAME, which is simply the inverse of DIFFERENT, a general faithfulness constraint. Any output that occupies a different point on the scale to the input violates this constraint. In this particular case, both [e] and [a] violate this constraint. The mechanism that blocks the A → C mapping is a constraint labelled ENDMOST. This constraint has exactly the same profile no matter what the input; for every step an output candidate is from the end of a scale, one violation mark is assigned. In this case, [a] does not violate ENDMOST. [e] is assigned one violation mark, and [i] is assigned two. The violation profile of ENDMOST is exactly the same no matter what the input. This is because it refers explicitly to how the scale is constructed, rather than any of its constituent parts.

In fact, all of the constraints that Mortensen uses to construct his model of how chain shifts should be represented are purely scale-referring, in contrast with, for example, Gnanadesikan's model. This is a necessary move for Mortensen, because he wishes to model scales that do not have phonetic substance behind them. This means, in effect, that the synchronic phonology can model any scale, regardless of how natural or otherwise that scale is.

However, this goes against much other work on how hierarchies in phonology work, for example de Lacy (2002; 2006) and Trommer (2009; 2011) (see chapter 5.4.2.2 for discussion). In these accounts, not only do scales have to have clear, phonetically interpretable trajectories, but (in Trommer's proposals at least) there are restrictions on the directions by which these scales can be traversed. On Mortensen's analysis there is no principled account of why we should get certain scalar patterns recurring time and again, or why the kinds of seemingly non-natural patterns that he discusses are rarer than these more common patterns.

## 2.4 Local Conjunction

A common thread that unites the three scalar approaches just discussed is that all require substantial changes to the architecture of the model in which they are couched. An approach to chain shifting that appears to be far less radical utilises Local Conjunction of constraints (Henceforth LC). LC has been a part of OT architecture since Smolensky (1993), and is defined thus:

- (11) Local Conjunction (Smolensky (1993)):  $C = [C_1 \ \& \ C_2]$  is violated iff both  $C_1$  and  $C_2$  are violated in a local domain  $D$ .

In essence this means that OT allows for the creation of constraints penalizing candidates that violate two specific constraints. LC has been used to model many and various phenomena (see Crowhurst (2011) for a recent review of the literature), and is the most common approach to chain shifting taken by OT theorists (see, for instance, Kirchner (1996), Moreton & Smolensky (2002), Walker (2005), Martínez-Gil (2006), Mascaró (2011). Łubowicz also discusses the method in various publications (2003a; 2011; 2012)). The essence of all of these approaches is the same. There will be some markedness constraint against input A, and some other constraint (or in some cases, the same constraint) against input B. These constraints exist independently, but can also be conjoined to form a highly ranked super-constraint that acts to rule out a mapping of  $A \rightarrow C$ , and force the  $A \rightarrow B$  mapping. If the constraint ruling out  $A \rightarrow A$  and  $B \rightarrow B$  is the same (see Moreton & Smolensky (2002) for examples of this), then the conjoined constraint is violated iff the original constraint is violated twice.

For a concrete application of this principle, we can turn to Martínez-Gil's (2006) analysis of the Lena Spanish shift using LC.

Table 2-5: Martínez-Gil's account of Lena Spanish raising (2006, p.134)<sup>7</sup>

/a/	*[-low, -ATR]	IDENT-[high] & IDENT-[low] <sub>seg</sub>	AGREE V-HEIGHT	IDENT-[high]	IDENT-[low]	IDENT-[ATR]
[i]		*!		*	*	*
[ɪ]	*!	*	*	*	*	
∅ [e]			*		*	*
[ɛ]	*!		*		*	
[a]			**!			

/e/	*[-low, -ATR]	IDENT-[high] & IDENT-[low] <sub>seg</sub>	AGREE V-HEIGHT	IDENT-[high]	IDENT-[low]	IDENT-[ATR]
∅ [i]					*	
[ɪ]	*!		*			*
[e]			*!			

The non-structure preserving candidates [ɛ] and [ɪ] are always ruled out by a highly-ranked constraint forbidding lax, non-low vowels (\*[-low, -ATR]). For the input /a/, the fully-faithful output [a] is ruled out by its multiple violations of AGREE V-HEIGHT. This constraint mandates, in this instance, that all vowels should have the same height value as the masculine suffix {-u}. In this interpretation of AGREE V-HEIGHT, the constraint appears to be gradient, which potentially allows the interpretation that scalarity is being let in through the back door. [a] and [e] both violate AGREE V-HEIGHT, but because [a] violates it more strongly, it is ruled out. However, there can be argued to be a principled phonological reason for [a] violating AGREE V-HEIGHT more times than [e]. Put simply, [a] disagrees with {-u} in the specifications of *both* of its height features ([+/- high] and [+/- low]), whilst [e] only differs from {-u} in its specification for [+/- high], as shown schematically in (12).

- (12) (a) /a/ [-high, + low]  
 (b) /e/ [-high, -low]  
 (c) /i/ [+high, -low]

The crucial part of this analysis is the conjoined constraint that rules out the candidate [i], IDENT (high) & IDENT(low)<sub>seg</sub>. This constraint is violated by any potential output candidate that differs from the input in terms of its values for both

<sup>7</sup> I have made certain simplifications to these tableaux. I have also made a correction in the second tableau. In my version, the [a] candidate violates AGREE V-HEIGHT twice, and in Martínez-Gil's version, [a] violates AGREE V-HEIGHT thrice. I believe that this is a mistake.

[+/- high] *and* [+/- low]. This constraint can also be accused of having a broadly similar motivation to scalar constraints, in this case with more justification. Indeed, this is explicitly acknowledged by theorists like Kirchner (1996). The function of the conjoined constraint is to allow a certain amount of deviation from the input without allowing too much. This, then, does not seem too different from Gnanadesikan's IDENT-ADJ constraint, or Mortensen's ENDMOST constraint, where structures that deviate from the input more than others incur more serious constraint violations, purely because of that deviation. Łubowicz (2011) draws attention to the similarity between the guiding influence behind Local Conjunction and scalar analyses (p.6). It is certainly true that Local Conjunction is a simple way of blocking an  $A \rightarrow C$  mapping, using architecture that has been part of the theory for many years, and that still has theoretical currency.

However, many phonologists<sup>8</sup> have raised the same questions about Local Conjunction: what may be conjoined, and what is an allowable domain of conjunction? Smolensky's definition does not specify whether there are limits on the kinds of constraints that can be conjoined. Nor does it specify the limits of the domain D. Subsequent attempts to refine this still leave a fairly open definition. For example, Moreton & Smolensky (2002) suggest that only certain kinds of constraints can be conjoined, ruling out combinations of MAX & Markedness, as well DEP & Markedness. However, Moreton & Smolensky give no real restrictions of what D can constitute. The authors note that others have made attempts to restrict the domain of chain shifting, and say "[t]he requirement discussed here, that the two constraints share a common domain, is less restrictive than any of these proposals, and hence compatible with them all" (p.2 [of online version]).

Despite the fact that there have been many proposals explicitly limiting the domain of conjunction (Moreton & Smolensky (2002) mention Crowhurst & Hewitt (1997), Fukazawa & Miglio (1998), Baković (1999), and Łubowicz (2002), to which I would add Łubowicz (2005)), they have not been adopted in any widespread fashion. The conception of LC as a mechanism with excessive computational power is pervasive and still discussed today (see, for example, Pater (2016)).

---

<sup>8</sup> See, for example, Ito & Mester (1998). McCarthy (1999), Padgett (2002), Łubowicz (2005)

Furthermore, in terms of its basic architecture, the mechanism of constraint conjunction seems to be as different from the way that constraints work in classical OT as the scalar theories discussed in the previous section. In classical OT, there are two kinds of constraints, markedness and faithfulness. A markedness constraint is violated by any candidate whose representation contains that marked structure, whilst a faithfulness constraint is violated by any candidate whose representation deviates from the input in the relevant respect. What causes a violation of a conjoined constraint is neither of these things, but a cumulative set of particular violations. The constraint itself requires two novel operations that the grammar does not otherwise need: the ability to copy constraints, so that one constraint can (in effect) be evaluated twice during EVAL; and the introduction of the logical operator &, which is not required for any other kind of constraint<sup>9</sup>. Whilst constraint conjunction has long been a part of OT, I contend that it introduces just as much baggage to the grammar as other, less well-known theories of chain shifting.

## 2.5 Contrast Preservation

Whilst all theories of chain shifting require some special mechanism to block an A → C mapping, the motivation behind this blocking is not always the same. In rule-based phonology, the rules one uses do not come equipped with a particular teleology. Upon seeing a description of a process in a language necessitating a counterfeeding order, a phonologist unversed in that language would need additional information to know why the counterfeeding order was active in the phonology. In traditional, rule-based phonology, it is impossible to include any such information in representations, though in later, autosegmental approaches, the operations of spreading and delinking are claimed to provide an explanation for many phonological processes (see e.g. Goldsmith (1976)). However, even in these later approaches, ordering relations are expressed in a purely mechanical way. We cannot know, just from looking at the representations, why the rules are ordered in the way that they are.

---

<sup>9</sup> Though see Wolf (2007a) for an interesting survey of possible logical connectives in OT, which comes out in favour of & being the only connective that could be plausibly introduced to the grammar.

The constraints blocking neutralization in scalar and Local Conjunction accounts are more explicitly teleological, which is a general property of Optimality Theoretic analyses. Indeed, as Hansson notes, “the output-oriented and constraint-prioritizing character of Optimality Theory means that phonological derivation...is construed as an inherently goal-oriented procedure” (2008, p. 868). In scalar accounts of chain shifting, there is a constraint that explicitly blocks movements that are ‘too far’ along some phonologically motivated scale. Similarly, locally conjoined constraints rule out excessive violations of constraints in a particular direction. Both of these theories, then, militate against outputs that are particularly unfaithful, with a side effect of this being the avoidance of a potential neutralization. However, the avoidance of neutralization, or its converse, preservation of input contrasts, is considered by Łubowicz (2003a, *et seq*) to be the key motivation behind chain shifts in both diachrony and synchrony.

Łubowicz’s Preservation of Contrast theory deviates from classical OT in two major ways. The first, and most radical, is that Łubowicz fundamentally alters the nature of GEN. Instead of individual candidates being evaluated against each other, Łubowicz uses the concept of ‘scenarios’. Table 2-6 is a set of potential scenarios for the Lena Spanish shift, based on the kinds of scenarios that we find in Łubowicz’s work (for simplicity’s sake I ignore the non-low back vowels, focusing only on vowel height):

*Table 2-6: Potential scenarios for the Lena Spanish shift*

Chain shift	Total Merger	Transparent	Identity
a → e	a → i	a → e	a → a
e → i	e → i	e → e	e → e
i → i	i → i	i → i	i → i

In Preservation of Contrast theory, it is the cumulative violations incurred by (in this case) the three mappings *as a whole* that is the key factor in which set of mappings is chosen as the optimal scenario. In a system based on contrast, it is not possible to judge candidates in isolation. A set of candidates that are very well formed in terms of markedness and/or faithfulness may well create a good deal of neutralization, but if the candidates are all evaluated separately, there is no systematic way to penalize

this. If contrast is to be central to how the grammar operates, the grammar must be able to evaluate optimal *systems*.

The second major innovation in the theory lies in the constraints themselves.

Łubowicz introduces a new family of constraints, named Preserve Contrast (PC) constraints, which are violated in cases of neutralization. There are three kinds of PC constraint: Input-oriented PC, Output-oriented PC, and Relational PC. Our discussion need only make reference to Input-oriented and Output-oriented PC constraints, but for detailed explanations of all three, see Łubowicz (2003a, pp.18-27).

Łubowicz discusses a shift in Kashubian Polish that is similar in character to that in Lena Spanish. What follows is based closely on this discussion (2003a, pp. 46-54). In an attempt to maintain coherence, I will try to show how Łubowicz’s analysis of Kashubian can be applied to Lena Spanish. The first important point is that Łubowicz treats the first step of the shift, which is /a/ → [e] in both Kashubian and Lena, as an instance of fronting, rather than raising. /a/ is specified as [+back] (p.47). Formally, this fronting is mandated by the highly-ranked markedness constraint \*CENTRAL. This constraint rules out what Łubowicz terms an ‘Identity’ scenario, in which all of the inputs are faithfully realized in the output (/A/ → [A], /B/ → [B], /C/ → [C]).

Table 2-7: *Lena in Contrast Preservation*

	*CENTRAL	PC <sub>IN</sub> (bk)	PC <sub>OUT</sub> (bk)	PC <sub>IN</sub> (high)	PC <sub>OUT</sub> (high)
☞ Chain Shift (a → e, e → i, i → i)				* {e,i}	* {i}
Transparent (a → e, e → e, i → i)		*! {a,e}	* {e}		
Neutralization (a → i, e → i, i → i)		*! {a,e}	* {i}	* {e,i}	* {i}
Identity (a → a, e → e, i → i)	*!				

It is at this point that Łubowicz’s PC constraints become crucial to the analysis. As can be seen in the tableau above, it is the PC constraints that rule out the ‘transparent’ scenario, where the A → B but not the B → C part of the shift occurs,

and a ‘neutralization’ scenario in which all inputs map to C. The tableau is based on Łubowicz’s Kashubian Polish tableau (p.51). However, I have added a neutralization scenario in which all inputs are realized in the output as [i].

Both the transparent and the neutralization scenarios violate the constraint  $PC_{IN}(\text{back})$ . Łubowicz defines  $PC_{IN}(P)$  constraints thus:

“For each pair of inputs contrasting in P that map onto some output in a scenario, assign a violation mark” (2003a, p.18)

In this specific situation, this essentially mandates that two inputs that contrast in backness in the input should also contrast in *some way* in the output, but that this surface contrast does not have to be one of backness. Thus, the chain shift scenario does not violate this constraint. /a/ and /e/ contrast in backness, and this contrast is preserved as a height contrast. Even though /e/ and /i/ no longer contrast in the output, both being realized as [i], they did not contrast for (back) in the input, so there is no violation of the constraint.

For the related Output-oriented constraint,  $PC_{OUT}(\text{back})$ , the violation profile is the same. Below is Łubowicz’s definition of Output-oriented constraints:

“For each output that corresponds to two or more inputs contrasting in P assign a violation mark” (2003a, p.20)

With regard to  $PC_{OUT}(\text{back})$ , this means that if two or more inputs that contrast in (back) are realized in the same way on the surface, the constraint is violated. Both the transparent and neutralization scenarios violate this constraint, because two inputs that contrast in backness are realized in the same way in the output. In the transparent scenario, /a/ and /e/ are both realized as [e], and in the neutralization, /a/ and /e/ are both realized as [i].

A crucial factor in Łubowicz’s analysis is that some contrasts are more important than others. Thus, the backness contrast is seen as the most important part of the shift. This motivates the general ranking schema  $PC(\text{back}) \gg PC(\text{high})$ . If this were

reversed, the transparent scenario would become optimal, because the chain shift scenario merges the input contrast on the height dimension (/e/ → [i], /i/ → [i]). Neutralization here, then, represents a ‘worst-of-the-worst’ scenario, because it merges contrasts with regard to both backness and height.

Contrast Preservation theory changes the architecture of the grammar in fairly radical ways. As stated at the start of this section, this is because of Łubowicz’s main conceptual argument: that even in synchronic grammar, phonological processes are sensitive to the effect that they have on the system as a whole. This touches on a fundamental question that has been at the heart of linguistics since at least the days of American Structuralism: to what extent should functional, teleological theories be a part of explaining how language works? The quote below, from Bloomfield, shows that there has long been a suspicion of teleological theories in linguistics:

“We have acquired understanding and the power of prediction and control and reaped vast benefit in the domains where we have developed non-animistic and non-teleologic science. We remain ignorant and helpless in the domains where we have failed to develop that kind of science” (1944, p. 55)

Even at this stage however, there was little agreement. Roman Jakobson, for example, took the polar opposite stance, stating, “language (and in particular its sound system) cannot be analyzed without taking into account the purpose which that system serves” (1928, p. 1). This presupposition is the foundation for theories of sound change such as those proposed by Martinet (1955), where function is the key motivation for theory. Chain shifts are a key set of examples in this case, at least in diachronic shifts, as after the original triggering movement, any subsequent movements occur to preserve the communicative function of the system. These movements can take the form of sounds moving away from other sounds that have begun to encroach on their phonetic space (commonly referred to, at least since Martinet (1955), as ‘push shifts’), or sounds moving into phonetic space that has recently been vacated by another sound (commonly referred to as ‘pull’ or ‘drag’ shifts).

However, this conception of the motivations for chain shifting is not unproblematic. The main issue is that, as Labov points out, functional arguments presuppose that large-scale merger of sounds should never occur. However, he argues, it unequivocally does (1994, p.551). For a simple example, consider Greek Iotacism, in which 8 separate vowels in Ancient Greek all merge, with a final realization, in all cases, of /i/ (Labov 1994, p.229). He also states that chain shift is less frequent than merger. This is based on his analyses of diachronic sound change and sound change in progress, but this can be tested on synchronic effects, at least to an extent, with reference to Gurevich's (2004) survey of lenition effects in synchrony.

It is true that the overall finding of Gurevich's study is that contrast avoidance is very common, and neutralization is rare. From her corpus of 230 lenition processes, she argues that only 18 processes (or 8% of all studied effects) result in genuine neutralization. However, chain shift is a very rare strategy of contrast avoidance. Whilst Gurevich does not provide the statistics in her breakdown, a study of her corpus reveals six lenition effects that could be described as chain shifts, a third of the amount of neutralizations. These shifts are listed in table 2-8<sup>10</sup>.

---

<sup>10</sup> These effects do not form part of my new corpus of chain shift effects in chapter 4, except for the Northern Corsican effect which is labelled Gran Canaria Spanish in the corpus. I discuss my general reasoning for that choice in the opening sections of that chapter.

Table 2-8: Potential chain shifts from Gurevich's lenition corpus

Language	Shift	Pages in Gurevich	Source(s)
Périgourdin French	$t \rightarrow d, d \rightarrow \delta / V\_V$	99-100	Marshall (1984)
Limbu	$p \rightarrow b, b \rightarrow w / V\_V$	145-146	van Driem (1987)
Malayalam	$p \ t \ k \rightarrow b \ d \ g,$ $b \ d \ g \rightarrow \beta \ \delta \ \gamma /$ [+son]_, [-nas]_V	153-154	Mohanan (1986)
Nepali	$ts^h \rightarrow h, h \rightarrow \emptyset / V\_V$	170-174	Bandu & Dahal (1971) Acharya (1991)
Northern Corsican	$p \ t \ k \rightarrow b \ d \ g,$ $b \ d \ g \rightarrow \beta \ \delta \ \gamma / V\_V$	178-180	Oftedal (1985), Dinnsen & Eckman (1977)
Senoufo	$p \ t \ k \rightarrow b \ d \ g,$ $b \ d \ g \rightarrow \beta \ \gamma \ / V\_V$	205-206	Mills (1984)

As stated previously, Łubowicz's model can deal with a wide range of scenarios, and also makes detailed predictions about the kinds of scenarios that should be logically possible. However, it stands or falls on one very specific assumption, i.e., that contrast transformation is the driving force behind not just some, but all genuinely phonological processes. Whilst the debate as to whether synchronic phonology can admit functional explanations is far from over, what is certain is that Preservation of Contrast theory restructures the entirety of the grammar. Thus, Łubowicz's claim that her theory provides a unified account of opaque and transparent processes with no additional machinery (a claim made, for example, in her thesis (2003a, p.vii)) is only true once a wholesale retooling of OT has been accomplished.

## 2.6 Optimality Theory with Candidate Chains

The previous sections illustrate either the most specific attempts to solve chain shift problems (scalar theories and contrast preservation), or the most widely known and used (counterfeeding orders, Local Conjunction). This final subsection discusses one of the most recent approaches, Optimality theory with Candidate Chains (OT-CC; McCarthy (2007)), which has been argued to account for chain shift effects of various kinds (see analyses in, e.g., McCarthy (2007), Nagle (2008), Gussenhoven & Jacobs (2011), Wolf (2011)).

The key defining feature of OT-CC is that, unlike in classical OT, derivation is necessarily serial. How fundamental a change this constitutes from classical OT is up for debate. Importantly, the creators of OT, Prince and Smolensky, did not mandate that processing had to proceed in parallel (1993/2008). Later analyses may have put parallel processing front and centre, but it is not a necessary design property of the model (see, e.g., Vaux (2008) for discussion). However, the opacity problem in OT is largely caused by the fact that, in the vast majority of later OT analyses, processing *is* necessarily parallel, meaning that the blocking of A → C mappings cannot result from the relative ordering of processes, as in SPE-style phonology.

In OT-CC, each candidate is represented via a ‘chain’; a series of one-step changes to the fully faithful candidate, all of which must lead to an output form that better satisfies the constraint ranking of the language under analysis than the previous step in the chain. A separate evaluation process occurs at each stage, until the most harmonic candidate possible is selected. The chain building operation is constrained by the design of GEN (see McCarthy (2007)), but also by a set of constraints introduced to the grammar by McCarthy called precedence, or PREC constraints. These constraints are violated when certain other constraints are violated in particular orders. For a concrete example, Nagle’s OT-CC analysis of a vowel raising chain shift in Bengali (2008) can easily be adapted to work for Lena Spanish. The tableaux below (adapted from Nagle (2008)) show how the OT-CC account works:

Table 2-9: OT-CC account of Lena (1)

/gatu/	*[+low]/_ {-u}	ID(low)	PREC(ID[high], ID[low])	*[-high]/_ {-u}	ID(high)
<gatu> FFC	*!				
☞ <gatu, getu> <ID(low)>		*	*	*	
<gatu, getu, gitu> <ID(low), ID(high)>		*	**!		*

The first line of each candidate shows the chain, with the rightmost form in the chain being the prospective output (in this case, [gatu], [getu], [gitu] from top to bottom). The bottom line shows the faithfulness constraints that are violated by each step. The

fully faithful candidate (in this case [gatu]) is ruled out by a high-ranking markedness constraint against /a/, as in every other analysis so far. However, the other two candidate chains <gata, getu> and <gata, getu, gitu> both violate ID(low), as the prospective surface forms are both [-low], whilst the input /a/ is [+low]. In this case it is the constraint PREC(ID[high], ID[low]) that breaks the tie. This particular precedence constraint can be violated in two separate ways.

- (13) (a) If a candidate violates ID(low) *before* it violates ID(high).  
 (b) If a violation of ID(high) occurs *after* a violation of ID[low].

Table 2-10 shows the result of the evaluation for the input /nenu/.

Table 2-10: OT-CC account of Lena (2)

/nenu/	*[+low]/_ {-u}	ID(low)	PREC(ID[high], ID[low])	*[-high]/_ {-u}	ID(high)
<nenu> FFC				*!	
☞ <nenu, ninu> <ID(high)>					*

The ID(high) constraint is violated before any violations of ID(low), because there are no violations of ID(low) at all. This means that it is the markedness violation incurred by <nenu>, of \*[-high] that causes the movement. It is worth considering just how similar this process is to the counterfeeding rule order discussed earlier in the section. Both the constraints \*[+low]/\_ {-u} and \*[-high]/\_ {-u} are analogous to rules mandating that /a/ and /e/ must both change before {-u}. The only difference is that the constraints do not specify what the target of the rule is. The ID constraints, their relative ranking, and the PREC constraint achieve this specification. ID(high) is a less serious violation, both in terms of its own ranking, and its primacy in the PREC constraint, than ID(low). The PREC constraint is a statement forcing the grammar to order changes in the specification of [+/-high] before changes in [+/-low]. The fatal violation in the first tableau is caused by a candidate chain in which the specification of [+/- high] is changed *after* the specification of [+/-low]. This is essentially analogous to creating two rules, in which B → C must occur before A → B. It is my

contention that OT-CC is simply an extrinsic counterfeeding order done with constraints.

The argument that OT-CC practitioners would make against this point, in favour of OT-CC constraints, is that the system is significantly more constrained in OT-CC than it is in rule-based phonology, where any arbitrary rule can be instantiated. However, there is no necessary limit to what constraints one may postulate either. In both cases, the analyses are made in the same way; the data is observed, by the linguist or by the learner, and representations are formed to suggest how those observations might be instantiated cognitively. In this respect, OT-CC is significantly more complex than rule-based analyses. Instead of two rules, five constraints are required, one of which has two separate conditions, both of which must be evaluated at each pass through the system. There is no additional conceptual information given in the OT-CC account either; the fact that the shift in Lena (and, at least on some accounts, the Bangla shift also<sup>11</sup>) is assimilatory in character is not hinted at by the constraints which model it. If one takes the position, as in scalar or Contrast Preservation theories, that the reason *why* a chain shift is happening should be fully interpretable from the representations that model it, then OT-CC is again equivalent to a counterfeeding account, in that the reasons for order of the PREC constraints are not made explicit.

## 2.7 Summary

This chapter, then, is a review of current approaches to synchronic chain shift. Whilst there are many theories that can model synchronic chain shift, all of them involve assumptions that are controversial to large sections of the phonological community. An argument that I have been making for some time (see, e.g., Neasom (2011; 2013)) is that a key reason for this lack of consensus is not simply theoretical affiliation. The problem is one of those discussed in chapter 1; that one author's chain shift is another's diachronic artefact, or set of processes randomly coming together, or set of predictable allomorphs. Before we can get back into detailed discussion of theories and their properties, we must begin to disentangle the definitional issues surrounding synchronic shifts.

---

<sup>11</sup> See, for example, Ghosh (2001)



# Chapter 3: Chain shifts: Synchrony vs. diachrony and acquisition

## 3.1 Introduction

As Łubowicz points out in her review article (2011), chain shifts have been reported in a wide variety of sub-disciplines in phonology, from dialectology to second language learning. In this section I discuss the two most commonly-discussed kinds of chain shift that are not synchronic effects in stable adult grammars: diachronic chain shifts and shifts in first-language acquisition. Analysis of these effects yields two findings: 1) there are theorists working on both historical and acquisition data who are using the same kinds of theories that we have seen applied to synchronic shifts; and 2) there are good reasons to believe that we should not actually do this. This section argues that there are few substantive similarities between chain shifts across domains.

## 3.2 Diachronic chain shift

The early study of chain shifting was entirely based on diachronic processes. Indeed, some of the most well-studied effects in historical linguistics can be, and frequently are, described as chain shifts. The classic example of chain shifting in terms of vowel quality is the Great English Vowel Shift (GEVS). In this shift, the long vowels of English changed completely in quality, with mid and low vowels raising and high vowels diphthongizing. This shift has been studied for well over a century (see, for example, Luick (1896) and Jespersen (1909)), and there is still active debate over the exact mechanics of the effect. As well as this, the Germanic consonant shifts described by Grimm's law and later by Verner's law have also been considered to be chain shifts (for a recent treatment in this vein, see Noske (2012)).

Since Martinet (1955), it has generally been considered that there are two kinds of diachronic chain shift; push shifts and pull, or drag, shifts (see chapter 2.5 for the definition of these terms).

I note at the outset that, in the case of both the GEVS and the Germanic consonant shifts, I have thus far been discussing what is essentially the textbook position: that both the GEVS and Grimm's law constitute coherent, totally interrelated push or

drag shifts. This is exemplified by, for example, Campbell (2004, pp. 46-48), or chapter 13 of Aitchison (2013). In both cases, though, and even in the cases of more recent shifts it is not certain that this is true. For instance, it is not possible to accurately reconstruct the time course of the Germanic consonant shifts that are influenced by Grimm and Verner's laws. Lass (1997, pp. 243-246) shows persuasively that there are many orders that the effects could potentially have applied in. As Honeybone points out, Lass' analysis "does not commit us to a chain shift analysis" (2001/2002, p. 53), because not all of these orders can be discussed as simple push or drag chains.

Moreover, Stockwell & Minkova (1988) and Stockwell (2002) argue that rather than one coherent, overall shift, there are various parts of the GEVS that interact rather less than is traditionally assumed. If it turns out that diachronic shifts cannot be distilled down to pushing or dragging, then there is even less reason to attempt to compare them to synchronic chain shifts like that in Lena Spanish, which for all intents and purposes do appear to be coherent effects where the  $A \rightarrow B$  and  $B \rightarrow C$  parts of the shift are reflexes of the same overall effect. However, for the rest of the section, I adopt something close to the textbook position, in order to keep the comparison between effects in synchrony and diachrony as direct as possible. Even making this simplifying assumption, however, I argue that there are substantive differences between the two sets of processes that go beyond the Saussurean (1916/2011) division between diachrony and synchrony.

### *3.2.1 How much should we be attempting to model?*

It is clear from the historical record that, at least in many cases, diachronic chain shifts take centuries to unfold. This is certainly the case with the GEVS, which is usually assumed to have operated from the 15<sup>th</sup> to the 18<sup>th</sup> century. This is potentially what leads Kirchner to assume that the representational issues that make synchronic chain shifting difficult are unproblematic when we consider diachronic chain shifts, "since the ordering of the sound changes presumably corresponds to distinct historical stages" (1996, p. 341). The implication of this is that for an individual speaker, the overarching chain shift has no psychological reality. Speakers acquire a set of underlying representations for the inventory of their language. All else being equal, these underlying representations will map faithfully to their surface

realizations. This is the default view in most frameworks, where it is widely agreed that phonological grammars do not recapitulate history for its own sake.

However, there are some approaches under which distinct historical stages can be modelled (see, e.g., Kiparsky (1973a), Bromberger & Halle (1989)). Indeed, Chomsky & Halle argue in *The Sound Pattern of English* (SPE, (1968)) that vowel shift in English is a productive, synchronic rule (see section 4.3 of SPE). Chomsky and Halle assume that in pairs like *div[ay]n – div[ɪ]nity*, the diphthong in the simple form is derived from underlying /i/, and that the fact that many pairs like this exist are the positive evidence required by learners of English to store underlying forms that are not identical to the surface form in words like *divine*. This, then, recapitulates the diphthongization that is present in the GEVS in the synchronic grammar.

McCawley strongly disputes the logic of this kind of argument. He argues that even though it is probably true that speakers make some kind of connection between related pairs of words, it is very unlikely that these connections are substantive enough to affect the storage of underlying forms (1986, p.28). Both he and Jaeger (1984) present experimental studies investigating the productivity of the English vowel shift rule, which Pierrehumbert sums up as evidence that the rule is “only partially productive” (2006, p.84). Indeed, Jaeger suggests that the ‘vowel shift rule’ is more likely to be a reflex of a generalization English speakers make based on orthography, rather than any principled grouping of vowels, a la Chomsky & Halle, stating that “there is no experimental evidence supporting the psychological existence of the specific vowel alternation group postulated by Chomsky & Halle” (1984, p.30).

In OT, there is also discussion about the extent to how the architecture of the theory should be used to discuss diachronic change. Holt (2003) is a collection of articles showing how OT might deal with diachronic processes, and includes an article by Miglio & Morén (2003) that discusses how the GEVS could be modelled in OT. The most recent account of this kind is presented by Fulcrand (2015), who criticises Miglio & Morén’s account of the GEVS because it is not, in his terms, unified. That is to say, Miglio & Morén postulate different constraint rankings for different

historical stages of the shift. Fulcrand presents one grammar that models all of the changes that occurred in the GEVS. His analysis uses the contrast preservation constraints proposed by Łubowicz (2003a; 2011; 2012) that I discuss in Chapter 2. Table 3-1 is the tableau taken wholesale from Fulcrand’s paper (2015, p. 568).

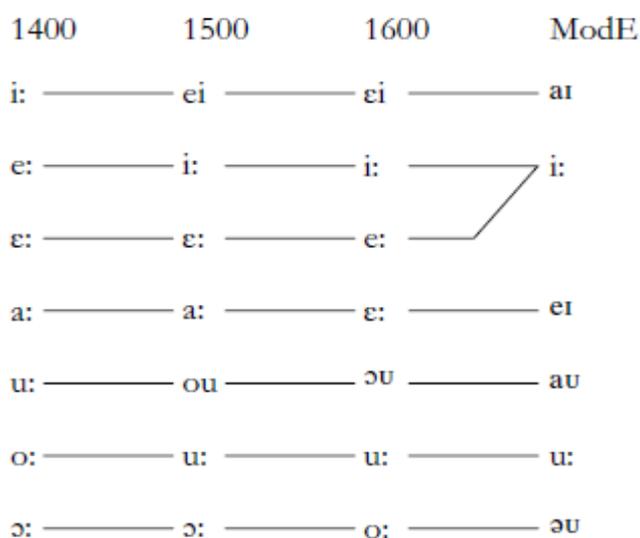
Table 3-1: Fulcrand’s analysis of the GEVS

/a:, ε:, e:, i:/	*V[+low]	PCIN(low)	PCIN(tns)	PCIN(hi)	*V[-tns]	INTEG
Identity [a: ε: e: i:]	*!				*	
Transparency [ε: ε: e: i:]		*!			**	
Transparency [a: e: e: i:]	*!		*			
Transparency [a: ε: i: i:]	*!			*	*	
Fusion [ai ai ai ai]		*!	*	*		****
☞ Opaque [ε: e: i: ai]					*	*

Recall that rather than evaluating individual candidates, Contrast Preservation theory evaluates over *scenarios*, or sets of mappings. In table 3-1, Fulcrand evaluates several potential results of the GEVS. The constraints are relatively straightforward. The markedness constraints \*V[+low] and \*V[-tns] assign violations to any members of a scenario that exhibit the marked feature ([a:] and [ε:] respectively on Fulcrand’s analysis). The three PCIN constraints state that if two segments in the input scenario contrast for the feature specified in the constraint, then they must also contrast for that feature in the output. For example, in the first transparent scenario, the contrast in the feature [low] between /a:/ and /ε:/ in the input is merged, as both surface as [-low] ([ε:]). Therefore, the constraint is violated. The low ranked INTEG constraint simply assigns a violation to any input monophthong that is realized as a diphthong in the output. In this tableau, then, Fulcrand models the entirety of the GEVS at a stroke. He concludes, “it has been shown that even though [Contrast Preservation theory] was initially designed to account for synchronic chain shifts, it can also explain diachronic chain shifts” (2015, p. 569). I contend, however, that this is a case where the fact that we *can* do something does not necessarily mean that we *should*.

The tableau that Fulcrand provides can, at most, be said to represent the final stage of the GEVS. By this final stage of the shift, it seems highly unlikely that the input scenario for a given speaker would actually be represented /a:, ε:, e:, i:/ anymore. Figure 3-1 shows the time course suggested by Lass for GEVS (1999, p. 72).

Figure 3-1: The approximate time course of the GEVS, according to Lass 1999



According to this time course, diphthongization and mid vowel raising had taken place by 1500. A speaker born in, say, 1580 would at most have limited evidence that, for instance, what they heard as [ei] had its origins in historical *i:*. The only cases in which we can plausibly suggest that there is positive evidence for this would be in alternations of the *divine/divinity* kind. However, the move from *i:* → *ai* is assumed to have spread across the entire class of such forms. This means that, in words where there are no such alternations, for example *cry, my, lie, fly* etc., the previous pronunciation would not be accessible to our 1580's speaker.

A more specific problem for Fulcrand's analysis is that the driver of the shift is the highly-ranked markedness of the low vowel /a:/<sup>12</sup>. If this constraint is not present in Fulcrand's tableau, then the scenario in which no shift at all occurs is the winner. However, figure 3-1 indicates that Middle English /a:/ remained /a:/ until the 16<sup>th</sup> century, meaning that it came after the changes to at least the diphthongs and high

<sup>12</sup> This is a consequence of using Lubowicz's system (2003a, 2012), where all shifts are kick-started by a high-ranked markedness constraint against the first element of the shift.

vowels. If this is true, then Fulcrand recasts the historical reasons for the actuation of the shift into something chronologically implausible. There is debate about whether the GEVS as a whole constitutes a push or a drag shift, or something else entirely. However, the motivation for the shift cannot be a change that happened centuries *after* other changes in the shift.

Lass eloquently argues that, when we make a historical statement of the sort  $x > y$ , “the juxtaposition of initial and final states is not the whole story; the mechanisms producing the  $t_1 > t_2$  transition are complex and often indirect” (1999, p. 77). This forms, to my mind, a sharp distinction between diachronic and synchronic shifts. In all analyses of synchronic chain shift that I am aware of, it is assumed that the *entire* chain shift is part of the synchronic grammar of the individual speaker.

### 3.2.2 *The roots of synchronic chain shifting*

As alluded to in the previous subsection, Łubowicz (2003a; 2011; 2012) does not insist on a division between shifts in synchrony and diachrony. This is presumably because the central organizing principle of her theory is the maintenance of contrast, which she argues is a plausible motivation for both diachronic and synchronic processes. In her review article, she makes the generalization that very often chain shifts involve vowel height. To show this, she gives the two examples in (1).

- (1) (a) New Zealand English (Labov, 1994):  $\text{æ} \rightarrow \text{e} \rightarrow \text{i} \rightarrow \text{i}$   
(b) Nzɛbi (Clements, 1991):  $\text{æ} \rightarrow \text{ɛ} \rightarrow \text{e} \rightarrow \text{i}$

The New Zealand English Shift (NZES) is a well-studied effect (see, e.g., Bauer (1979), Trudgill et al. (1998), Gordon et al. (2004), Langstrof (2006)) that has taken place since at least the 1850s. The short front non-high vowels have risen and the short front high vowel has centralized. This has been described as both a push and a pull chain, though the modern consensus appears to be in the direction of a push shift (see, e.g., Gordon et al (2004) and Langstrof (2006)). The Nzɛbi effect (originally described by Guthrie (1968), and subsequently studied by Clements (1991) and Kirchner (1996)) is argued to be a genuinely synchronic chain shift. In all analyses that I am aware of, the effect in Nzɛbi is taken to be a morphologically triggered phonological process, analogous to the Lena Spanish effect discussed in previous

chapters. Certain affixes, usually containing the suffix {-i}, condition a one-step raising effect. For example, the root form /sal/, when affixed to the verbal suffix {-i}, is realized as [sɛl-i] ‘to work’. Simultaneously, /βɛɛd + i/ is realized as [βeedi] ‘to arrive’ (examples from Kirchner (1996, p.344)).

On the surface, then, the two shifts do not appear to be too dissimilar, leaving aside their different termination points. However, there are several factors that suggest that the similarities between shifts like those in New Zealand English and those in Nzebi are at best superficial. I have spoken about the Nzebi case up to this point because it is the example that Łubowicz uses in her article. However, historical information on the effect is scant, so in what follows I refer to similar effects in other languages whose history has been discussed more explicitly. My definition of similar is that, as in Nzebi, the shifts I go on to describe have all been analysed as regular synchronic processes that are triggered by certain classes of affixes.

Modern analyses of the NZES (see, e.g., Gordon et al. (2004) and Langstrof (2006)) suggest that the effect is a gradual push shift that began with words in the TRAP vowel class<sup>13</sup>. These words encroached into the phonetic space of words in the DRESS vowel class (*bad* becoming more like *bed*, for example). In turn, DRESS vowels changed quality, becoming more like words in the KIT class (*bed* becoming more like *bid*). Finally, words in the KIT class centralized. Gordon et al. are fairly clear that this is the order in which these changes occurred, and that the movements constitute distinct stages:

“[T]here is no significant degree of DRESS raising without also TRAP being raised, and there are very few speakers with KIT centralization who do not also have raised DRESS, but there are speakers who have close TRAP but not close DRESS and speakers with close DRESS who do not have centralized KIT” (Gordon, et al., 2004, p. 266).

Given that different speakers have different surface forms, it does not seem unreasonable to assume that there are differences in the grammars of these speakers.

---

<sup>13</sup> Here, like the authors referenced, I use Wells’ lexical sets to describe categories of words that share the same vowel in a particular accent of English (1982)

In a sense, it is unimportant whether that is at the level of representation (i.e., whether speakers reanalyse their underlying forms from generation to generation) or at the level of computation (i.e., some speakers have rules in their grammar that other speakers do not).

It is also worth noting that Gordon et al. are considering the *overall* values for vowel quality for sets of words across speakers. In a gross sense, the shift can be said to be unconditioned, in that it is generally assumed that the end result of the shift<sup>14</sup> is a reorganization of the vowel system of the language. Whilst Langstrof's analysis in particular highlights the fact that the changes of the shift take place with slightly different effects, and with a greater or lesser frequency, depending on various phonological and sociolinguistic factors (2006, pp. 157-159), there is an overall change in the vowel quality in the vast majority of environments.

This can be contrasted with a shift like those in Nzebi or Lena Spanish. I have found no plausible examples of synchronic chain shifts that are unconditioned (section 5.2.3 touches on this briefly). However, one could argue that this is perhaps not so surprising, and instead is just a function of how diachronic and synchronic analyses work. An unconditioned synchronic chain shift would result in widespread absolute neutralizations. Consider a bizarro-world version of Lena Spanish, where the raisings we observe are completely unconditioned. The result of this would be a language without surface [a], because *every* underlying /a/ would necessarily be raised to [e] by virtue of the raising rule. In the SPE-style literature, this kind of absolute neutralization was possible. However, this kind of reasoning is now largely seen as unnecessarily abstract (a particularly stinging critique of absolute neutralization can be found in Lass (1984, pp. 211-213)). Therefore, the straightforward analysis of bizarro-world Lena would be that there is no /a/ in the grammars of modern speakers of that language, even if historically we can see that there used to be /a/, but a raising rule eliminated it.

Indeed, a logical consequence of this might be to assume that synchronic chain shifts are remnants of diachronic effects that persist in a sub-part of the grammar,

---

<sup>14</sup> It would perhaps be more accurate to say 'current state' here, given that language change is never really over.

potentially as a result of morphologization. Indeed, this is how Maiden describes metaphony effects (1991), and the explanation that Schmidt (1996) gives for a vowel raising effect in the Cameroonian language Basaa, which is similar in many ways to the Nzɛbi effect<sup>15</sup>. However, the historical processes whose residues have become synchronic chain shifts are not diachronic chain shifts like the GEVS or NZES.

Even with a relatively large amount of relatively high quality corpus data, it is not certain what initiated the changes in the New Zealand English shift. Gordon et al. suggest that fronting of the START vowel may have been the catalyst for the movement of the TRAP vowel, though they are cautious about this because the START vowel is a long vowel. This would indicate that a change in one subsystem of vowels can affect another subsystem, which seems unusual in the light of the other shifts discussed in, for example, Labov (1994). The trigger is much more obvious in the case of metaphony effects and the Nzɛbi effect, and is completely different. Maiden's (1991) analysis of metaphony effects in Italian dialects envisages historical metaphony as "an essentially scalar, dynamic, unidirectional assimilatory process" (p.111) where all vowels assimilate towards following high vowels. Schmidt discusses the Basaa raising shift as "a remnant of the cross-height vowel harmony inherited from proto-Volta-Congo through the Benue Congo branch" (1996, p. 242). Essentially, the raising process in Basaa, which can be schematized as /a/ → [e], /ɛ/ → [e], /e/ → [i], recapitulates a [+/- ATR] harmony process in proto-Bantu.

Whilst in both of these cases the synchronic chain shifts appear to be effects that are relics of historical changes to vowel height, those changes were assimilations in specific contexts, not the kind of reorganizations of the vowel system that we would traditionally associate with chain shifting. Even under a more detailed typology of conditioned and unconditioned changes like that in Honeybone (2001/2002), these would genuinely be "conditioned" changes, which can be thought in some way to be directly 'caused' by the environment in which they occur...*(clear examples are assimilations and dissimilations)*" (p.25, emphasis mine).

---

<sup>15</sup> Schmidt attempts to include some of this history in her account, which does include absolute neutralizations. For other analyses of Basaa that do not require absolute neutralization, see Parkinson (1996) or Jesney (2005).

Vowel height chain shifts in synchrony appear to have little to do with vowel height chain shifts in diachrony, and more to do with these kinds of assimilation processes. This may go some way towards explaining certain typological properties of synchronic chain shifting, which would be puzzling if we had to draw a direct parallel between diachronic and synchronic vowel height shifts. For example, many diachronic chain shifts involve fronting or centralizing as well as shifts in height. In my sample of synchronic shifts (see chapter 4), there are no instances of changes in backness in any of the vowel shifts. This is easily explained if they are the results of harmony processes that are uninterested in the specification of [+/- back].

In sum, I would argue that the only links between diachronic and synchronic chain shift appear to be their  $A \rightarrow B \rightarrow C$  patterning. Given that there are substantive differences when we look a little closer, the conservative position is to assume that the processes are not related in a principled way. In the next section, I discuss shifts in L1 acquisition, where the assumption that there are substantive similarities with synchronic chain shifts in adult grammars is even more widespread.

### 3.3 Shifts in L1 acquisition

Chain shifts have been part of the conversation on generative approaches to child language acquisition since Smith's extensive diary study of his son Amahl's phonological acquisition (1973). This book contained discussion of the now famous *puzzle – puddle – pickle* shift. In this shift, Amahl would realize underlying fricatives as stops, so a word like *puzzle* would be realized as [pʌdl]. However, a velarization rule, turning coronal stops into velars before /l/ was also in effect. Therefore, the word *puddle*, assumed to be underlyingly /pʌdl/, thus identical to Amahl's surface form for *puzzle*, would be realized as [pʌgl]. We can schematize the Lena shift and the *puzzle – puddle – pickle* shift in the same way.

- (2) (a) Lena Spanish:  $a \rightarrow e \rightarrow i$   
(b) Amahl:  $z \rightarrow d \rightarrow g$

A priori, there appears to be no reason why we should not at least contemplate using similar grammatical mechanisms to model chain shifts in L1 acquisition and in

stable adult grammars. Indeed, Smith discusses the possibility that, in order to model the *puzzle – puddle – pickle* shift, it may be necessary to use extrinsic rule ordering, though it should be noted that he argues against this position in later work on the grounds that extrinsic orders are unlearnable (see Smith (2010, p. 2), where he cites Miller & Chomsky (1963, p. 430)). However, it is true that extrinsic counterfeeding rule orders can model chain shifts in both acquisition and adult grammars, and that they can do so in the same way.

In terms of Optimality Theoretic approaches, there appears to be a reasonably widespread assumption that there are substantive similarities between shifts in acquisition and shifts in stable adult grammars. Take, for example, McCarthy's (2005) article on free rides, which analyses a putative shift in the stable grammar of Sanskrit (this is discussed at length in chapter 5, section 3). The other example of a synchronic chain shift that McCarthy gives comes from Dinnsen & Barlow's (1998) paper on child language acquisition. That paper discusses a shift that can be schematized as  $s \rightarrow \theta \rightarrow f$ . This shift is also mentioned by Mortensen (2006), when he states that his scalar method of modelling chain shifts is capable of modelling shifts in acquisition. Researchers in clinical phonology have couched treatments of chain shift effects in the speech of children with phonological disorders in the same terms as theories on adult chain shifts. For example, Dinnsen et al. (2011) use OT-CC to model the  $s \rightarrow \theta \rightarrow f$  shift. In all of these cases, it is assumed that the basic workings of both child and adult chain shifts are essentially the same.

Similarly, Jesney (2005; 2007) presents an OT treatment of child chain shifts, based on what she terms 'faithfulness to input prominence'. In essence, Jesney's theory is based on the idea that certain features are particularly perceptually salient. Even when inputs that contain these features are not permissible outputs in the child's grammar, the particularly salient feature is retained. (see Jesney (2005, pp. 76-89) for her analysis of the *puzzle – puddle – pickle* shift). Jesney tentatively suggests that her theory could be profitably applied to synchronic chain shifts in the adult grammar, giving an example of how it could work in a vowel raising shift in Basaa (see chapter 4 for more on this shift, and Jesney (2005, p. 164) for exposition of this point). My question in this section, then, is not whether it is possible to represent shifts in child

grammars and adult grammars in the same way, as it has been repeatedly shown that it is. The more interesting question is, again, whether we *should* do this.

### 3.3.1 *The reality of child chain shifts: Evidence from diary studies*

Given that child chain shifts do not, in general, persist into adulthood, it is worth thinking in more detail about what actually occurs during the ‘chain shift’ stage of development. I acknowledge that chain shifts can be stubborn problems for children who exhibit them as part of a speech disorder (see, e.g., Dinnsen et al. (2011)). However, in children who do not have such disorders, any chain shift will usually be in evidence for a fairly small portion of their phonological development. It has even been argued that, at least in some cases, the appearance of a seeming chain shift can take on a greater significance than perhaps it warrants. I present two examples below of investigations into how significant certain child chain shifts really are.

Ettlinger (2009) presents a diary study of “a typically developing [American-English learning] child, M, between the ages of 1;0 and 2;0” (p.2). The child appeared to produce a chain shift that was the result of an interaction between velarization and stopping in word-initial position ( $s \rightarrow t, t \rightarrow k$ . The examples in (3) come from Ettlinger (2009, p. 3)).

- (3) (a) [tak] ‘sock’  
(b) [kʌk] ‘talk’

On one hand, M is incapable of making the [s] sound required for the adult pronunciation of the word *sock*, which they realize with a [t]. On the other hand, M *is* capable of making the [t] sound required for the production of the word *talk*, but in this context they do not, instead velarizing to [k]. Ettlinger argues that the appearance of a chain shift here is in fact something of a phantasm. By the time the chain shift should have been active in the child’s grammar, the velarization rule had already stopped productively applying to new forms. Consider table 3-2, showing the ‘chain shift’ stage of M’s development.

Table 3-2: Ettlenger's illusory chain shift (taken from 2009, p.4)

	Stage 1 (1;0 – 1;4)	Stage 2 (1;4- 1;7)
(a) <i>cookie</i>	kʊ.ki	kʊ.ki
(b) <i>talk</i>	kʌk	kʌk
(c) <i>sock</i>	n/a	tak
(d) <i>table</i>	n/a	tej.bo

The words *sock* and *table* are not produced by M in the first stage of the diary study, whereas the words *cookie* and *talk* are. If *table* is indeed a word that is not simply missed by the diary study, but actually only *acquired* in stage 2 (which Ettlenger claims is the case (2009, p. 4)), then this is troubling for the chain shift account. In sum, there is a stage where velarization is occurring (Stage 1), and we have no evidence to suggest that stopping is in operation. In the next stage, stopping is clearly operating, as we can see from the newly acquired *sock*, but the fact that *talk* still surfaces with velar [k] is argued by Ettlenger to simply be a holdover from the previous generalization of velarization, which is no longer active. He calls this effect 'lexical inertia'. A term from Menn and Stoel-Gammon (1993). If the two processes are acting at different times, then the effect is in no way comparable to any kind of synchronic chain shift.

Another case where an apparent chain shift does not actually seem to represent a stage in the child's development that we would wish to instantiate a grammar for comes from Smith (2010). In his Grandson Zachary's speech (henceforth Z), Smith notices a putative shift that he refers to as the 'horse-shoe' shift, because Z's pronunciation of the term 'horse-shoe' gives an illustration of the effect of the shift in microcosm. Z's production of the term is [hɔ:t su:]; alveolar /s/ surfaces as [t] word-finally, whilst palato-alveolar /ʃ/ surfaces as [s] in all positions. This creates a word-final shift of ʃ → s → t. As the 'horse-shoe' case indicates, there appear to be times at which the effects of both processes can be seen on the surface.

As Smith points out, though, the environmental restrictions on the two rules are quite different. Smith divides Z's development into stages, usually lasting between 1 and 2 months. The first stage where the rule becomes salient is his stage 7 (when Z was aged between 2;5;15 and 2;7). In his description of the phonology of this stage, Smith notes that "/ʃ/ was [s] initially, but became [t] elsewhere" whilst "/s/ was usually [s] initially...but was often [t] finally" (p.70). This suggests two

neutralizations, /s, ʃ/ → [s] word-initially and /s, ʃ/ → [t] word-finally, rather than a chain shift.

It turns out that this is the stage of development whence the ‘horse-shoe’ example comes. This example was recorded when Z was 2 years, 6 months and 21 days old, towards the end of Smith’s stage 7. During this specific stage of development, Smith notes that, word-finally, there is variation between [s] and [t] for underlying /s/. On the other hand, “/ʃ/ is [s] initially, [d] medially, and [t] finally” (p.81). This is borne out by a search of Smith’s diachronic lexicon, which shows 7 instances of final [t] for target /ʃ/ and no instances of [s] for target /ʃ/ during the period of time that Smith refers to as stage 7 of Z’s development. That is to say, a ‘horse-shoe’ shift was not in effect at the time that Z uttered the phrase. The stopping rule affects both /s/ and /ʃ/ in the same way word-finally, and does not affect word-initial /s/.

The subsequent stage of Z’s development (Smith’s stage 8, when Z was aged between 2;7;4 and 2;9;7) is the last stage in which the stopping rule can be said to be in fairly regular operation. However, it does not appear that there is a systematic chain shift occurring, according to Smith’s descriptions. At stage 8, /s/ and /ʃ/ both exhibit similar variation word-finally. Both can be realized as [t, d] or [s]. Table 3-3 shows how common these rates of application are at this stage, according to the data in Z’s diachronic lexicon:

*Table 3-3: Relative rates of application of the variants for word-final /s/ and /ʃ/ in Z’s stage 8*

	/s/		/sh/	
	[s]	[t,d]	[s]	[t,d]
Occurrences	85 (64.4%)	47 (35.6%)	15 (68.2%)	7 (31.8%)

Whilst the quantities of word-final /s/ and /ʃ/ are very different, the proportions of surface fricatives and stops are roughly the same in both cases<sup>16</sup>. The pattern in stage 8 suggests that what we are really witnessing is the death of the stopping rule. This

<sup>16</sup> This reflects the relative prevalence of final /s/ and /ʃ/ in the lexicon. A search of Lindsey & Szigetvari’s online pronunciation dictionary, CUBE (<http://seas3.elte.hu/cube/>) gives 11,122 entries for final [s] and just 552 entries for final [ʃ].

change cannot be expected to be totally uniform and regular, particularly since many of the words on which this data is based appear only once in Z's diachronic lexicon. Stage 8 lasts 65 days. It is the only stage in which we observe a significant number of forms that can conceivably be thought of as showing a chain shift. In stage 7, the stopping rule is predominant in both word-final /s/ and /ʃ/. By Stage 9, whilst the /ʃ/ → [s] rule is very much in operation, the /s/ → [t] rule is almost completely dead: “/t, d, l, n, s, z/ are systematically correct in all positions” (2010, p. 90). It is worth asking whether it is plausible that the learner could place a constraint enforcing the chain shift into the grammar for such a short period of time, before removing it so completely that it never surfaces again.

Consider, for example, a Local Conjunction account of the horse-shoe shift (for a Local Conjunction account of the s → θ → f shift, see Morrisette & Gierut (2008). For accounts arguing that LC is not the method we should be using, see Jesney (2005, 2007) and Walton (2012)). A constraint like IDENT[+/- anterior] & IDENT[+/- cont]<sub>seg</sub>, would do the required work of ruling out a segment that changed its specification for [+/-anterior] *and* [+/-continuant], whilst allowing violations of the individual constraints to have less of an impact. Table 3-4 shows how this would work.

Table 3-4: Tableaux of a Local Conjunction analysis of the horse-shoe shift

/ʃ/	*ʃ	IDENT[+/-ant]&[+/-cont] <sub>seg</sub>	IDENT[+/- ant]	IDENT[+/-cont]
ʃ	*!			
☞s			*	
t		*!	*	*

/s/	*ʃ	IDENT[+/-ant]&[+/-cont] <sub>seg</sub>	IDENT[+/- ant]	IDENT[+/-cont]
ʃ	*!		*	
☞s				
t				*!

The real issue at stake here is whether this is a desirable move. Inserting a constraint like this to deal with a recalcitrant effect that is not found in adult grammars feels like an ad hoc reaction to a potential problem, rather than a move with any deeper justification. In sum, one potential explanation for the horse-shoe effect in Z's productions is that it took some time for the stopping rule to completely die out (for potential reasons for this, from a classical OT perspective, see Ettliger (2009)), and that there was some degree of overlap because the /ʃ/ → [s] rule was still active. The alternative is to state that, for around two months, Z had some sort of constraint in his phonological grammar that had all of the properties in (4).

- (4) (a) *Optionality*: In many cases, the chain shift mappings failed to occur even during the stage where both were most prevalent.
- (b) *Novelty*: Until stage 8, the mapping for word-final /ʃ/ was [t]. Until that point, there was no evidence for an  $A \rightarrow C$  ( $f \rightarrow t$ ) blocking mechanism. Post stage 9, there is no need to postulate such a mechanism. This is because the reason usually given for such contrivances is that /B/ (in this case /s/) maps unfaithfully. Post stage 9, /B/ maps completely faithfully, meaning that there is no reason to suppose /A/ would map anywhere else.
- (c) *Limited shelf-life*: After the ~65 days that the constraint was active, it went from being completely undominated to having no power whatsoever, and remained in this state for the rest of Z's phonological development.

I do not believe that this kind of constraint is plausible. Though I use Local Conjunction in my example, I believe that this extends to any rule or constraint based analysis that employs an  $A \rightarrow C$  blocking mechanism.

### 3.3.2 *More general concerns*

In recent years, there has been a growth in the literature on chain shifts in acquisition that seeks to position such effects as problems of performance, rather than competence. This would accord with the findings of the previous section, in which I suggested that using the grammar to model marginal effects like the horse-shoe shift was a somewhat brute-force solution to a set of circumstances which are not necessarily even a problem. Smith (2010) is clear that “[m]ost of the rules or constraints of child phonology are motivated by performance considerations rather

than competence ones” (p.43). He points out that this is a position that has been argued for strongly by Hale & Reiss (2008), who explicitly place problems like child chain shifts outside of grammatical competence. Walton (2012) studies the *puddle – pickle* and  $s \rightarrow \theta \rightarrow f$  shifts and concludes that the grammar is not well-equipped to model these effects, suggesting instead that they represent performance errors in both articulation and perception. Below I discuss some of the reasons the authors adduce for these claims.

From the perspective of this thesis, though, it is first worth considering the fact that certain effects in child chain shifting would be very difficult to motivate in the stable adult grammar. Consider the  $s \rightarrow \theta \rightarrow f$  shift that I discuss above. A large-scale study on this shift was conducted by Dinnsen et al (2011), where they found that it recurred across many English-acquiring children (example data from Dinnsen & Barlow 1998, p.66):

- (5) *Subject 33 (Age 5;4)*
- $/s/ \rightarrow [\theta]$
- (a)  $[\theta i?]$  ‘sink’
- (b)  $[\theta \Lambda n]$  ‘sun’
- (c)  $[d\epsilon\theta i]$  ‘dress’ (dim)
- (d)  $[a i \theta \Gamma]$  ‘icy’
- (e)  $[v e i \theta]$  ‘vase’
- (f)  $[d u \theta]$  ‘juice’
- $/\theta/ \rightarrow [f]$
- (g)  $[f \Lambda m]$  ‘thumb’
- (h)  $[f o r m]$  ‘thorn’
- (i)  $[t i f i]$  ‘teeth’ (dim)
- (j)  $[m a u f i]$  ‘mouth’ (dim)
- (k)  $[b \ae f]$  ‘bath’
- (l)  $[t u f]$  ‘tooth’

In examples (5a-5f), underlying  $/s/$  is realized as  $[\theta]$ , and in examples (5g-5l), underlying  $/\theta/$  is realized as  $[f]$ . Underlying  $/f/$  is realized faithfully (see Dinnsen &

Barlow (1998, p. 66) for examples). The examples show that both of the mappings are effectively unconditioned. The chain shift appears to occur in word-initial, word-medial, and word-final position. It should be noted that the effect is not unconditioned in all of the children who participated in the study (see Dinnsen et al., (2011, p. 2)). However, the important generalization is that in child chain shifts it is at least possible for one or both parts of the shift to be unconditioned. As discussed in the previous subsection, it is unlikely that synchronic chain shifts work in this way.

What, then, is the basis for a chain shift analysis of these problems? A child with the  $s \rightarrow \theta \rightarrow f$  shift does not have an [s] sound in their surface inventory at all. As well as this, the only instances of  $\theta$  in the child's grammar come from what is assumed to be underlying /s/. However, the extent to which we should assume that the child's system of phonology is similar to an adult's is uncertain. Whilst, as we have seen, some acquisitionists do take the position that we can model processes in acquisition using the same grammatical techniques that we use to model adult speech, there are dissenting voices who suggest that we should not imbue language acquiring children with an identical set of grammatical tools to an adult (see especially Hale & Reiss (2008), Walton (2012)).

There is the argument, first put forward by Macken (1980), that the reason for Amahl's *puzzle* – *puddle* – *pickle* shift is that the rules are fundamentally different in nature. The stopping rule (*puzzle* → *puddle*) operates exceptionlessly, and appears to genuinely be a rule based on Amahl's articulations (Macken terms the process an 'output rule'). However, Macken argues that the velarization rule is grounded in perception (she terms the rule a 'perceptual encoding' rule). In her analysis, adult pronunciations pass through a perceptual filter before they enter the child's lexicon, in essence shaping the child's underlying forms. Rules like stopping then act on these lexical representations. This gives us an account of the shift that is not at all mysterious. There is no need, for instance, for extrinsic ordering. The rules could not be ordered in any other way, because perceptual and output rules operate in different domains. This is one of the few arguments that I do carry over into my discussion of chain shifting in synchronic adult grammars, though the domains in question are different (see Chapter 5, section 5.4).

Also, if we take the standard position that a chain shift is an  $A \rightarrow B \rightarrow C$  pattern where the mappings are expected to be from underlying to surface forms, we can argue that there is no  $A \rightarrow B \rightarrow C$  mapping here, only  $A \rightarrow B$ . Macken notes that the perceptual filter does not always create URs like /pʌgəl/. Unlike stopping, which is exceptionless in A's grammar, Macken shows that there are many counterexamples to the *puddle* – *puggle* generalization (10/47 tokens, or 21.3% of her sample). In these 10 cases, Amahl stored an adult-like /d/ form and produced it. In the other cases, he stored something that sounded more like /g/. The only way to form a coherent chain shift here is to suggest that the first part of the chain shift is not Amahl's underlying form, but actually the adult production. If we take this position, then we can see an  $A \rightarrow B \rightarrow C$  order. Adult |z|  $\rightarrow$  [d] in Amahl's grammar, and adult |d|  $\rightarrow$  [d] or [g]. This kind of transmission of mappings is not how chain shifts in adult grammars work, as it is always assumed that the entirety of the chain shift is present in the grammar of the speaker.

Hale & Reiss (1998; 2008) dispute the notion that Smith's *puzzle* – *puddle* – *pickle* shift is a genuinely phonological problem at all. They argue that, rather than a phonological mapping relationship from /z/  $\rightarrow$  [d], and another from /d/  $\rightarrow$  [g] / \_\_1, what we actually hear from A are performance errors which end up in close approximations of [d] and [g] respectively. If the chain shift is really just a function of a stage of development where a language learning child has not yet mastered adult-like pronunciations, then two things are true. Firstly, such processes are not really chain shifts, in the sense of an  $A \rightarrow B \rightarrow C$  mapping. Underlying /A/ is not really mapping to something that is identical to underlying [B], as is assumed to be the case in synchronic chain shifts. In terms of a mapping relation, what we really have here is more like  $A \rightarrow A'$ ,  $B \rightarrow B'$ ,  $C \rightarrow C$  (see Walton (2012, p. 67) for elaboration of this point). Secondly, such processes are not the same as those that we observe in adult grammars, where, if we follow the assumptions of the great majority of the literature, the mappings that we observe do meet this  $A \rightarrow B \rightarrow C$  criterion. The most explicit statement of this line of argumentation in the literature comes from Walton (2012), who postulates that the *puzzle-puddle-pickle* shift and the  $s \rightarrow \theta \rightarrow f$  shift are motivated not only by performance concerns, but by the same performance concerns in both cases. She integrates Macken's perceptual explanations with Hale & Reiss' more articulatory explanations.

“The first error pattern, [θ] for /s/ and *puddle* for *puzzle* words, is due to articulatory difficulties with /s/ and /z/. The second error pattern, [f] for /θ/ and *puddle* for *pickle* words, is due to misperception of /θ/ and /dl/.” (p.103)

On Walton’s account, then, stopping (which occurs in both shifts) is an articulatory problem where speakers are unable to produce strident fricatives. Whilst the processes that comprise the B → C parts of the *puzzle – puddle – pickle* shift and the s → θ → f shift are different, Walton presents detailed arguments that both of these problems are perceptual in nature (see her chapter 3, section 3.4). On accounts of this nature, there is no accurate comparison that can be drawn between apparent chain shifts in acquisition and stable adult grammars.

Walton (2012, p.67) also avers that the s → θ → f and the *puzzle – puddle – pickle* shifts should not be considered chain shifts because they are made up of two independent processes. It is certainly true that there are no child chain shifts that work in the same way as an adult shift like the Lena Spanish shift, where it seems to be clear that both the A → B and B → C parts of the shift are intrinsically related. However, relatedness of process is not a sine qua non of chain shifting in the adult grammar, at least under certain theories (see, e.g., Moreton & Smolensky (2002), Łubowicz (2003a)).

Moreover, Dinnsen et al. argue that we cannot be entirely sure that the processes in the s → θ → f shift *are* independent, given that the processes do not appear to be free to interact in every logically possible way. In Dinnsen et al’s study, some of the 160 children studied did exhibit the chain shift pattern (11 participants, or 6.875% of the sample). However, the most common pattern is for speakers to have only half of the shift. That is to say, fully 59% of the sample (94 participants) exhibited only the θ → f part of the shift, pronouncing underlying /s/ faithfully as [s]. A much smaller group, of three participants (2% of the study’s population), exhibited only the s → θ part of the shift. This shows that both parts of the shift can occur completely independently in the grammar of any particular child. Dinnsen et al point out that no child in the study exhibits a neutralization, in which /s/ and /θ/ are realized as [f]. However, if we again consider the argument made most forcefully by Hale & Reiss (2008) and Walton (2012), that mappings such as these are not moderated by the grammar but

are instead problems of performance, then we get a more principled explanation for why the only effect that we appear to get when children exhibit both error patterns should be something that looks like a chain shift.

Even if we allow adult-like underlying representations for /s/ and /θ/, the sounds that these URs map *to* are not necessarily identical to the underlying form of the adjacent segment. That is to say, /s/ is not mapping on to [θ], just a sound that approximates to it. Walton (2012) discusses this point in detail, pointing to several studies that suggest that children acquiring language have difficulty producing /s/ (e.g., Snow (1963), Velleman (1988) Smit (1993), Gibbon (1999), Morrisette & Gierut (2008)). These studies also suggest that its mispronunciations for /s/ cannot always be described as [θ], even if the difference between fronted /s/ and /θ/ cannot be perceived without instrumental analysis (see especially Velleman (1988)). This is similar to the position taken by Bleile in his *Manual of Articulation and Phonological Disorders*:

“During lisping [s] and [z] are produced with the client’s tongue either touching the teeth or protruding slightly between the teeth as for [θ] and [ð]; however, the tongue may be either more forward or more retracted than for [θ] or [ð].” (2004, p. 150).

Lisped productions, then, may approximate [θ] or [ð], but equally may not. It appears, then, that in cases where a child attempts to produce /s/ and actually produces something that closely approximates adult [θ], the fact of this close approximation is actually accidental, and is not motivated by a grammatical rule or constraint.

Finally, Dinnsen et al. note that we do not observe ‘grandfather effects’, which can be defined by the mapping relation  $A \rightarrow C, B \rightarrow B, C \rightarrow C$  ( $s \rightarrow f, \theta \rightarrow \theta, f \rightarrow f$ ). If we follow Walton’s assumption that dentalization and labialization are distinct error patterns that should not be expected to interact, the lack of both neutralizations and grandfather effects is completely unproblematic. In fact, if we wish to postulate any kind of competence based account we would need to observe neutralization and/or grandfather effects, because they would offer evidence that the output of one error

pattern ( $s \rightarrow \theta$ ) can act as the potential input for another ( $\theta \rightarrow f$ ). By contrast, there are three mapping relationships we would expect on a performance-based account, and we get all of them:

- 1) None of the error patterns are in evidence.
- 2) One or the other error pattern is in evidence and the other is not.
- 3) Both of the error patterns are in evidence (the supposed chain shift pattern).

In conclusion, I concur with Hale & Reiss, Smith (2010) and Walton that it is more productive to see chain shifts in acquisition as performance errors, be that in the perceptual or the articulatory systems (or both). The main consequence of this is that it does not appear to be a valid move to compare shifts in acquisition with shifts in stable adult grammars.

### **3.4 Summary**

In this section I have presented arguments showing that the connections between chain shifts in synchrony and other domains are not as principled as has sometimes been assumed. In the first part of the chapter I argued that it is highly unlikely that diachronic chain shifts are mentally represented in a way that is genuinely similar to most models of synchronic chain shifts, as it seems implausible that hundreds of years of phonological development are recapitulated in the synchronic grammar. Also, I have illustrated that, at least in the case of common vowel raising shifts, the kinds of shifts that we observe in the synchronic grammar are remnants not of push or drag chain shifts, but of historical assimilation or harmony processes.

I have presented two main arguments against the existence of chain shifts in child phonology. First, I argue that it is often difficult to tell if a genuine chain shift is in operation due to the rate that children develop. Seeming chain shifts may represent, at best, very short stages of development, which I argue do not constitute justification for the introduction of a genuinely phonological rule or constraint. Second, there are methods of modelling child chain shifts that do not treat these kinds of effects as being genuine problems of competence, which make compelling predictions about the kinds of interactions we should expect between error patterns in acquisition.

Overall, then, this chapter presents justification for the narrowing of focus adopted in the rest of the thesis. I believe that it is worthwhile to make a firm separation between seeming chain shift effects in acquisition, diachrony and synchrony.



# Chapter 4: The limits of chain shifting, and a new chain shift corpus

## 4.1 Introduction

In the previous chapter, I attempted to disentangle the idea of a chain shift in synchrony from effects that have also been called chain shifts in diachrony and acquisition. Now that this distinction has been made, it is possible to turn our attention to those shifts that have been argued to be genuinely synchronic. From this point onwards I will only be discussing context-sensitive processes in which both the  $A \rightarrow B$  and  $B \rightarrow C$  movements have been argued to both be synchronically active in the grammar of one individual speaker of the language or dialect, at the same time and across the board.

The study of putative shifts in synchrony has been the focus of my previous research (Neasom 2011; 2013), and this thesis can be seen as a continuation of this line of enquiry. Indeed, the inspiration for this entire project is the same compendium of putative synchronic chain shifts, created first by Moreton & Smolensky (2002) and later updated by Moreton (2004a), which I focused on in Neasom (2011).

This chapter is, first and foremost, an expansion and reanalysis of the work started by Moreton, and the bulk of the chapter will be a presentation of an updated chain shift corpus, which I call the Revised Corpus of Putative Synchronic Chain Shifts. There are two dimensions of expansion. First, I provide more biographical information about each of the shifts that are listed in Moreton's version of the chain shift compendium. The second dimension of expansion is simpler, and constitutes the addition of 19 further examples of putative shifts from the wider literature<sup>17</sup>. The next section discusses these points in detail, describing Moreton's compendium and explaining and justifying the differences between that compendium and the sample in this thesis. I also give some details of the other major sources that I have used in constructing the sample, as well as a brief discussion of how the sample is presented. Section 3 is the sample itself, whilst section 4 presents a quantitative examination of

---

<sup>17</sup> I do not claim that the sample I have taken is completely exhaustive, but I do contend that it is the largest, most detailed sample that is currently available.

trends within the sample. Section 5 is a brief conclusion, introducing the notion that a great many examples of putative shifts in the corpus appear to be somewhat problematic.

## 4.2 The state of play

### 4.2.1 Moreton & Smolensky (2002) and Moreton (2004a)

As discussed in chapter 2, Moreton & Smolensky (2002) provide an analysis of chain shifts based on conjoined constraints. As well as this, they provide a selection of 35 putative chain shift processes in an appendix. Realizing the utility of this resource, Elliott Moreton subsequently posted a version of this ‘compendium’ on his website (Moreton 2004a), and has made significant additions. Currently, the compendium contains 49 putative shifts.

It is vital to note from the outset that, as mentioned in chapter 2, Moreton & Smolensky take the widest possible view of what can constitute a synchronic chain shift. This, from the front matter of the online version of the compendium, is the most explicit statement of Moreton’s position:

“This is a collection of known or alleged instances of the phonological phenomenon variously known as:

- “counterfeeding”
- “synchronic chain shift”
- “underapplication”” (Moreton 2004a)

Given the sceptical nature of this thesis, it would be disingenuous not to point out that Moreton certainly shares some of this scepticism. In the quote above, for instance, it is significant that the effects are ‘known **or alleged**’ instances of chain shifting. Additionally, whilst Moreton’s compendium has been cited uncritically as the go-to collection of chain shifts by Jesney (2005) and Łubowicz (2011), Moreton’s own aims seem to be somewhat more modest. The original version of the corpus was designed simply to test the prediction that there were no chain shifts involving combined violations of MAX & MARKEDNESS or MAX & DEP, which Moreton & Smolensky’s version of LC would rule out.

The compendium takes the same basic form in its original and newer incarnations, and a screen-grab from the online version is shown in figure 4-1.

Figure 4-1: Screen-grab from Moreton's corpus

Language	Mapping	Sources
Chemchuevi	$V_1V_2\# \rightarrow V_1\# \rightarrow \#$	Press 1979 UC Pub Ling 92, §1.33.
Chukchee	$\#CCV \rightarrow \#^{\cdot}CCV \rightarrow \#^{\cdot}CC^{\cdot}$ [underlying $\cdot$ can trigger harmony]	Spencer 1999 citing Bogoraz 1922
Egyptian Arabic	$V:h\# \rightarrow V:\# \rightarrow V\#$	Michael Becker p.c. 2002
English (1)	$\begin{matrix} h & h & h \\ p & t & k \end{matrix} \rightarrow p t k \rightarrow b d g$	
English (2)	$alt^{\cdot} \rightarrow aIR^{\cdot} \rightarrow a^{\cdot}R^{\cdot}$	Thomas 2000 <i>JPhonet.</i>
English (3)	$nIs \rightarrow ns \rightarrow nts$ (doubtful)	Donegan & Stampe 1979 Dinnsen (ed.), <i>Current Approaches to Phonol. Theory</i> ,
English (4)	$ald \rightarrow a^{\cdot}t \rightarrow alt$	B. L. Anderson (1999), NWA 28.
Etxarri Navarrese Basque (similar processes in other W. Basque varieties)	$e \rightarrow i \rightarrow i^{\cdot}$ $o \rightarrow u \rightarrow u^{\cdot}$	Kirchner 1995, citing de Rijk 1970 <i>Fontes Lg. Vasconum</i> 2(5):149-167
Finnish	$pp\ tt\ kk \rightarrow p\ t\ k \rightarrow v\ t\ \emptyset$ (complex environment)	Jensen & Stong-Jensen in Koutsoudas 1976 (ed.) <i>Application and Ordering...</i>
Hellendoorn Dutch	$ktn \rightarrow kn \rightarrow kN$	van Oostendorp 2002 "Morphological recoverability"
Hidatsa	$V_1V_2\# \rightarrow V_1\# \rightarrow \#$	Kenstowicz & Kisseberth 1977:178f. citing Harris 1942
Icelandic	$aCr\# \rightarrow aCur\# \rightarrow \ddot{o}Cur\#$	Kenstowicz 1994:79-81.
Irish	$ptk \rightarrow fhx \rightarrow \emptyset h?$ (RotB for $x \rightarrow ?$ )	Gnanadesikan 1997:189ff. citing Ní Chiosáin 1991 diss.

Putative shifts are arranged by language, their mapping is shown schematically, with at least one reference per shift, and there is a full bibliography at the end of the corpus. There is little discussion in either Moreton & Smolensky (2002) or Moreton (2004a) about the selection process involved. Indeed, it seems to have been very informal. Moreton (2004a) describes the process thus: "Some [shifts] are taken from the secondary literature; others were found by flipping through reference grammars in the library looking for telltale phrases like "This rule does not apply to long vowels created by Rule 27"" (front matter). The chain shift corpus that forms the majority of this chapter is based primarily on these 49 putative shifts, all but two of which are re-examined.

The first shift that I have not re-examined is a putative shift in Egyptian Arabic that is schematically rendered as  $V:h\# \rightarrow V:\# \rightarrow V\#$ . The source given for this shift is a personal communication from Michael Becker. I contacted Dr Becker and he

informed me that the source was Ellen Broselow's doctoral thesis, which is not currently in print. I have contacted Professor Broselow but at the time of writing she has not provided me with the details of the process. The second un-reexamined shift comes from English, taking the schematic form  $aIt' \rightarrow aIR' \rightarrow \check{a} R'$ . I have left this shift out of the corpus because, despite a close reading of E. Thomas (2000), the source that is cited in Moreton's compendium, I do not understand the schematic given in the corpus well enough to confidently discuss it.

#### *4.2.2 Other sources and entry criteria*

As well as re-examination, a goal of the new chain shift corpus that I have constructed is to widen the sample of synchronic chain shifts. However, it is important to limit this widening, so that the corpus does not become unduly speculative. For this reason, I have introduced the criterion that any effect introduced into the corpus must have been explicitly claimed in at least one study to be a synchronic chain shift. There are many effects that appear to have the same superficial  $A \rightarrow B \rightarrow C$  form as a chain shift in the literature, but I do not wish to ascribe a chain shift analysis to a pattern which the original authors did not describe as such. Therefore, for instance, whilst Naomi Gurevich's lenition sample (2004) includes five effects that could be seen as chain shift patterns (see chapter 2), I do not include them in my corpus because she does not label them as such. An additional example comes from Trommer (2011), who discusses similar processes in the Nilotic languages Thok Reel and Mayak. Though both processes have an  $A \rightarrow B \rightarrow C$  mapping pattern, Trommer is explicit that he considers the Thok Reel effect to be a chain shift and the Mayak effect not to be. Therefore, only the Thok Reel effect is included in the corpus.

There are two additional sources that also contain surveys of certain kinds of chain shift effects, all of which I do include in my new corpus. David Mortensen's dissertation (2006), discussed in chapter 2, presents not only a new theory of how chain shifts should be modelled, but also a wide range of data from tone shifts in Min and Hmongic languages. In the majority of the chain shift literature, discussion of tone shifts is limited to discussion of whether circular tone shifts should be considered genuine synchronic processes, usually with specific reference to a tone circle in the Southern Min dialect of Xiamen (see, for instance, Moreton (2004b),

Zhang, Lai, and Sailor (2006), Barrie (2006), G. Thomas (2008), Łubowicz (2011)). I have included the Xiamen shift in my corpus for completeness, given that Mortensen, at least, considers it a synchronic process that requires a synchronic explanation. Barrie (2006), and (more tentatively) Łubowicz (2011) offer support for a synchronic treatment of the Xiamen tone circle, though it is worth pointing out that Zhang, Lai and Sailor (2006) wug-tested the productivity of the circle, and found that as a whole, speakers did not apply the complete pattern to novel forms (see chapter 6 for further discussion of this point).

Tone circles aside, Mortensen's dissertation also presents a range of effects (one of which is discussed in chapter 2) in which tone sandhi effects take a form that is more similar to the canonical  $A \rightarrow B \rightarrow C$  pattern characteristic of the rest of the putative shifts in the sample. These are included as well, as there is no a priori reason to suppose that these effects are any less a chain shift than any of the other processes in the corpus. I should point out at this stage that I do not include every example of a putative shift cited in Mortensen's dissertation in my sample. This is because there are many dialects which Mortensen claims show evidence of chain shift effects, but he does not give examples or any kind of detailed discussion. Additionally, the sources cited for these effects are often not readily available, making it difficult to independently confirm Mortensen's assertions.

As will become more apparent throughout the thesis, a kind of process that appears to be one of the most 'canonical' kinds of shift involves changes in vowel height. Many such processes are discussed in Parkinson's dissertation (1996), which is a study of partial height-harmony processes. In partial height-harmony, a harmonic trigger causes assimilation, but the target does not necessarily take on the exact value of the trigger. For example, a hypothetical underlying form /pat + i/ would be realized as [peti] as opposed to \*[piti]. This  $a \rightarrow e / \_ i$  pattern should be familiar from my constant referencing of the Lena Spanish effect in chapter 2. Therefore, a number of partial height-harmony processes constitute putative chain shifts, and Parkinson's sample is a useful additional resource.

#### 4.2.3 *The Revised Corpus of Putative Synchronic Chain Shifts: A user's manual*

The corpus that follows in the next section is, like Moreton's corpus, arranged alphabetically by language. Similarly, it includes a schematic mapping of the shift, as well as a list of references. In terms of the shifts that are included in Moreton's corpus, I have sometimes added references that clarify the pattern, or offer updated analyses, whether or not they conflict with the original analysis. From this point, however, the corpus departs from Moreton's compendium, offering a range of additional information. Below, I list the features of each corpus entry.

**Language (dialect):** The language in which the shift occurs. If the shift only occurs in a specific dialect, this is listed in parentheses. Only one spelling of each language name is listed. For entries that are featured in Moreton's corpus, I use his spelling, unless it conflicts with that used by authors of primary literature on that language. For entries that are not featured in Moreton's corpus, I use the spelling found in the primary literature.

*Example:* Spanish (Lena)

**ISO code:** A three-character code that uniquely identifies the language. The codes used in the corpus are taken from that language's entry on the online version of ethnologue (see [www.ethnologue.com/codes](http://www.ethnologue.com/codes) for more information). The code is included to facilitate searching, either within the corpus or online, particularly as many languages have names that can be spelt multiple ways.

*Example:* spa<sup>18</sup>

**Language family:** The family to which the language belongs. This is again taken from the ethnologue entry for that language, and the entire familial path is traced (as in Gurevich's lenition corpus). Only giving the overall language family is insufficiently detailed, but as different languages have trees of different lengths, there is no principled or consistent way of picking a second or third branch to add more detail.

*Example:* Indo-European, Italic, Romance, Italo-Western, Western, Gallo-Iberian, Ibero-Romance, West Iberian, Castilian

---

<sup>18</sup> Note that ISO codes do not reliably distinguish between dialects, so shifts affecting different dialects of the same language will not have separate ISO codes.

**Spoken in:** Taking a useful idea from Lavoie's (1996) corpus of lenition trajectories, a key influence on Gurevich (2004), this section lists the specific countries and, if the language is sufficiently localized, regions in which the language is spoken. This information is either taken from the ethnologue entry for that language, or if this is not available, as in the case of some lesser known dialects, from primary literature discussing the language or dialect.

*Example:* Lena principality, Asturias region, Spain

**Process(es) involved:** This section is subdivided into 'A → B' and 'B → C' sections, and gives a two to three word description of the processes involved. I have attempted to be as general as possible, in order to make trend-spotting easier. For instance, there are 7 lenition trajectories present in the corpus, but lenition can be realized in multiple ways. Whilst I do note the specific trajectory taken in parentheses, the overall description given to all of these processes is 'Consonant lenition'.

*Example:* A → B: Vowel Raising

B → C: Vowel Raising

**Domain:** A description of the phonological domain over which the shift takes place. Moreton & Smolensky (2002) briefly mention three kinds of domain; featural (changes to features at the level of the segment), segmental (complete loss or addition of a segment) and prosodic (involving suprasegmental features). It is worth noting that they do not apply these labels to individual shifts, and as such there are some shifts in the corpus that fall beyond the purview of these labels. Two shifts from Moreton's own version of the corpus show that these labels are insufficient:

(1)  $iC_{(0)}l \rightarrow iC_{(0)}l \rightarrow iC_{(0)}\lambda$ : In this shift, from the Barrow dialect of Inupiaq (as discussed in Kaplan (1981)), there are two changes which would, under Moreton & Smolensky's definition, have to be considered to be 'featural'. In the A → B part of the shift, /i/ fronts to [i]. In the B → C part of the shift, /l/ palatalizes to become [λ]. I argue that if this were classified as a 'featural' shift, it would imply a false commonality between shifts like this, where multiple segments are affected, and shifts like Lena Spanish, where only vowels in a certain position can be affected.

(2)  $a \rightarrow i \rightarrow \emptyset$ : In this shift from Bedouin Hijazi Arabic (see Al-Mozainy (1981), McCarthy (1999) or Almiḥmadhi (2011)), the  $A \rightarrow B$  part is clearly featural in nature,  $/a/ \rightarrow [i]$ . However, in the  $B \rightarrow C$  part of shift, the change must surely be segmental, as  $/i/$  is deleted completely. A simple solution would be to allow for mixed categories such as ‘featural/segmental’, but I feel that this dilutes the terms to the point where they have little meaning of their own.

Because of this, I use my own system of domain labelling, with four kinds of shifts represented in (3).

- (3) (a) *Isomorphic*: The domain of the rule changing  $A \rightarrow B$  *completely overlaps* with the domain of the rule changing  $B \rightarrow C$ . In practice, this is very similar to Moreton & Smolensky’s ‘featural’ category, in that shifts labelled isomorphic are those in which only one segment changes. Note that this definition would exclude the Barrow Inupiaq effect in (1), but include the Bedouin Hijazi Arabic effect in (2).
- (b) *Adjacent*: The segment affected by the  $A \rightarrow B$  process is adjacent to the segment affected by the  $B \rightarrow C$  process.
- (c) *Overlapping*: The segment affected by the  $A \rightarrow B$  process is not adjacent to the segment affected by the  $B \rightarrow C$  process, but the environments in which each of the processes occur do have some shared ground.
- (d) *Non-overlapping*: There is no apparent connection between the  $A \rightarrow B$  and  $B \rightarrow C$  processes, either in terms of segmental adjacency or the environments of the two processes.

These labels have been chosen as they are sufficiently general to cover almost all of the shifts represented in the sample. In Chapter 5, I discuss those shifts which do not fit into these labelling criteria.

**Schematic:** This entry corresponds to the ‘mapping’ column in Moreton’s compendium. It shows how the shift operates, using an  $A \rightarrow B \rightarrow C$  template as its base. This is the entry in which we see the true nature of the shift. If possible, these schematics are taken wholesale from Moreton’s compendium, in order to give a consistent representation for the same shift. I have marked all schematics that are copied directly from Moreton’s compendium with (M). Moreton’s corpus does not use the IPA, instead relying on a SAMPA-like system. In my version of the corpus, IPA is used throughout. All other schematics are my own, generally because no schematic form is given in the literature discussing the process.

*Example:*  $a \rightarrow e \rightarrow i$

**Environment(s):** This entry gives a description, in standard  $A \rightarrow B / C\_D$  notation, of the environment in which the shift occurs. If only one description is given, it can be assumed that both the  $A \rightarrow B$  and  $B \rightarrow C$  processes occur in the same environment; if this is not the case, then the environments for both processes are given. Because, in many cases, there are separate environments for both processes, I list these environments separately to the overall shift schematic, in order to preserve the coherence of these schematics. A slightly unfortunate corollary of this decision, coupled with my decision to faithfully reproduce Moreton’s schematics where possible, is that some of Moreton’s schematics include environmental information (for much more discussion of this, and the problems that it causes, see the next chapter). This means that, in some cases, the environment information that is given is redundant.

*Example:*  $/\_ \{-u\}$

**Source(s):** All potentially useful sources are listed. This includes primary and secondary sources, updating the entries in Moreton’s corpus where necessary.

*Example:* Hualde (1989), Kaze (1989), Parkinson (1996), Gnanadesikan (1997), Puente (1996), Martínez-Gil (2006), Walker (2005)

**Examples:** Examples of both parts of the shift are listed here. The goal is to give two or three examples of each part of the shift. However, this is not always possible as the data for some of the effects is scant<sup>19</sup>. The exact source for the examples is listed for easy reference.

*Example:* A → B: [gata] ‘cat’ (fem. sg.) vs. [getu] ‘cat’ (masc. sg.)  
[santa] ‘saint’ (fem.sg.) vs. [sentu] ‘saint’ (masc. sg.)  
[blanka] ‘white’ (fem. sg.) vs. [blenku] ‘white’ (masc. sg.)

B → C: [nena] ‘child’ (fem. sg.) vs. [ninu] ‘child’ (masc.sg.)  
[kordero] ‘lamb’ vs. [kordiru] ‘lamb’  
[seka] ‘dry’ vs. [siku] ‘dry’  
(*All examples from Gnanadesikan (1997, p.209)*)

**Description:** This section briefly discusses how each shift works mechanically. In some cases, this will be fairly apparent from the foregoing information, but in other cases the processes involved require additional discussion. This section also gives a brief description of the kinds of analysis that have been applied to the shift in the literature, if any.

The new corpus, then, is intended to give additional detail to Moreton’s original examples and to expand the sample to include shifts that have not yet been addressed by Moreton. Given that a primary use for this corpus is as a resource for anyone interested in synchronic chain shift, I attempt to refrain from any commentary in the corpus about what I consider a synchronic chain shift to be.

---

<sup>19</sup> The scantness of this data is, to my mind, a cause for real concern. See the following chapter for discussion of this issue.

### 4.3 The Revised Corpus of Putative Synchronic Chain Shifts

**Language:** A-Hmao (Eastern dialect)

**ISO Code:** hmd

**Language Family:** Hmong-Mien, Hmongic, Chuangqiandian

**Spoken in:** Southwestern China (various provinces)

**Process(es) Involved:**  $A \rightarrow B$ : Tone Sandhi  
 $B \rightarrow C$ : Tone Sandhi

**Domain:** Isomorphic

**Schematic:**  $H \rightarrow M \rightarrow L$

**Environment(s):** /{H,MH}\_\_

**Source(s):** Mortensen (2006)

**Examples:** *All examples from Mortensen (2006, p.78)*

$A \rightarrow B$ :  $H \rightarrow M$   
 $tu^H$  ‘son’ +  $ki^H$  ‘grandchild’  $\rightarrow$   $tu^H ki^M$  ‘descendants’  
 $lfi^{MH}$  ‘long time’ +  $nti^H$  ‘long’  $\rightarrow$   $lfi^{MH} nti^M$  ‘for a long time’

$B \rightarrow C$ :  $M \rightarrow L$   
 $qu^H$  ‘old’ +  $t\text{sho}^M$  ‘clothing’  $\rightarrow$   $qu^H t\text{sho}^L$  ‘old clothing’  
 $dzfi^{MH}$  ‘animal’ +  $mpa^M$  ‘pig’  $\rightarrow$   $dzfi^{MH} mpa^L$  ‘beast of burden’

**Description:** In A-Hmao, high tones (H) become mid (M) following either a high or mid-high contour tone. This environment does not constitute a phonological trigger, but Mortensen asserts that in an earlier form of the language (which still has a modern analogue in Dananshan Hmong, also listed in this corpus) the trigger *was* phonological but subsequent changes have obscured this: “the substance of the tones conditioning the sandhi alternations has changed, but their behaviour relative to the other tones has not” (2006, p.83). As with all of the tonal effects addressed in his thesis, Mortensen models this effect with scalar OT constraints of his own devising (see chapter 2.3).

**Language:** Arabic (Bduul dialect)

**ISO Code:** avl

**Language Family:** Afro-Asiatic, Semitic, Central, South, Arabic

**Spoken in:** Jordan (Petra)

**Process(es) involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Stress Assignment

**Domain:** Non-overlapping

**Schematic:** #VCC#  $\rightarrow$  #'VCiC#  $\rightarrow$  #VC'iC# (M)

**Environment(s):**  $A \rightarrow B$ : /C\_\_C,  $B \rightarrow C$ : Final syllable

**Source(s):** Bani-Yasin & Owens (1984), Moreton (2004a), Neasom (2011)

**Examples:** *All examples from Bani-Yasin & Owens (1984, pp.209-210)*

$A \rightarrow B$ : #VCC#  $\rightarrow$  #'VCiC#

/ism/  $\rightarrow$  [ísim] 'name' (Bani-Yasin & Owens, 1984, p. 210)

/jibt/  $\rightarrow$  [jíbít] 'I bore' (Bani-Yasin & Owens, 1984, p. 209)

$B \rightarrow C$ : #'VCiC#  $\rightarrow$  #VC'iC#

/ganam/  $\rightarrow$  [ganám] 'sheep' (Bani-Yasin & Owens, 1984, p. 210)

/bugar/  $\rightarrow$  [bugár] 'cattle' (Bani-Yasin & Owens, 1984, p. 210)

**Description:** In the Bduul dialect of Arabic final stress in CVCV(C) forms falls on the final syllable. The standard stress assignment is that which is listed in the C forms [ganám] and [bugár]. However, underlying #VCC# sequences undergo a process of epenthesis, creating #VCiC sequences. This epenthetic [i] cannot take stress, meaning that stress falls on the penult, as in [ísim]. Patterns such as this have been widely discussed in the literature (see, for example, Broselow (2008) or Alderete (2001) for OT accounts), but rarely as examples of chain shifts, even in work that discusses both chain shifts and stress-epenthesis processes (e.g., Łubowicz 2003a). A discussion of problems with classifying the Bduul Arabic pattern as a chain shift can be found in Neasom (2011).

**Language:** Arabic (Bedouin Hijazi dialect) 1

**ISO Code:** acw

**Language Family:** Afro-Asiatic, Semitic, Central, South, Arabic

**Spoken in:** Saudi Arabia (Hejaz Region), Eritrea

**Process(es) involved:**  $A \rightarrow B$ : Vowel Reduction (Vowel Raising)  
 $B \rightarrow C$ : Vowel Reduction (Vowel Deletion)

**Domain:** Isomorphic

**Schematic:**  $a \rightarrow i \rightarrow \emptyset$  (M)

**Environment:** Unstressed syllables

**Source(s):** Al-Mozainy (1981), McCarthy (1993), Orgun (1995), Kirchner (1995), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *All examples from Kirchner (1995, p.2)*

$A \rightarrow B$ :  $a \rightarrow i$

/katab/  $\rightarrow$  [kitab] ‘he wrote’

/samiʕ/  $\rightarrow$  [simiʕ] ‘he heard’

/rafaagah/  $\rightarrow$  [rifaagah] ‘companions’

$B \rightarrow C$ :  $i \rightarrow \emptyset$

/ʕarif-at/  $\rightarrow$  [ʕarfat] ‘she knew’

/kitil/  $\rightarrow$  [ktil] ‘he was killed’

/kitil-at/  $\rightarrow$  [kitlat] ‘she was killed’

**Description:** In the first part of this shift, unstressed /a/ raises to [i]. As Bedouin Hijazi Arabic has a three vowel system, /i a u/, this is a one-step vowel raising process. In the second part, the same environment, unstressed position, causes underlying /i/ to be deleted entirely. In the three Optimality-Theoretic accounts of the effect (McCarthy (1993), Orgun (1995), and Kirchner (1995)), it is assumed that the “shift constitutes a unified phenomenon of vowel reduction” (1995, p. 2). Kirchner uses the Bedouin Hijazi Arabic data as an early test case for a Local Conjunction account of synchronic chain shift (1995, p.10).

**Language:** Arabic (Bedouin Hijazi dialect) 2

**ISO Code:** acw

**Language Family:** Afro-Asiatic, Semitic, Central, South, Arabic

**Spoken in:** Saudi Arabia (Hejaz region), Eritrea

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Vocalization)  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Overlapping

**Schematic:** badw  $\rightarrow$  badu  $\rightarrow$  bidu (example from Al-Mozainy (1981), reproduced in McCarthy (1999, p.8)) (M)

**Environment(s):**  $A \rightarrow B$ : /C \_\_,  $B \rightarrow C$ : / \_\_.C

**Source(s):** Al-Mozainy (1981), McCarthy (1999), (2002), (2007), Moreton & Smolensky (2002), Moreton (2004a), Almihamdi (2011)

**Examples:** See schematic, above

**Description:** In the Hijazi dialect of Bedouin Arabic, a process of vocalization affects underlying word-final glides, such as that in /badw/, which becomes [badu]. However, the language also has a process in which /a/ raises to become [i] in open syllables. For example, “underlying /nasi/ 'he forgot' surfaces as [nisi]” (McCarthy, 1999, p. 14). The expectation is that /badw/ would undergo both processes and surface as [bidu], but this does not happen. It is worth noting that Moreton's schematic suggests that /badu/  $\rightarrow$  [bidu] is a representation of a genuine lexical item in BHA, but there is no suggestion in McCarthy's work that this is the case. McCarthy proposes three different Optimality Theoretic explanations; Sympathy Theory (1999), Comparative Markedness (2002), and Optimality Theory with Candidate Chains (2007). However, it should be noted that McCarthy himself does not describe the process as a chain shift.

**Language:** Arabic (Bedouin Hijazi dialect) 3

**ISO Code:** acw

**Language Family:** Afro-Asiatic, Semitic, Central, South, Arabic

**Spoken in:** Saudi Arabia (Hejaz region), Eritrea

**Process(es) Involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Non-overlapping

**Schematic:**  $\text{gabr} \rightarrow \text{gabur} \rightarrow \text{gibr}$  (M. Note: *this schematic is problematic. See discussion below*)

**Environment(s):**  $A \rightarrow B$ : /C\_\_C,  $B \rightarrow C$ : /\_\_.C

**Source(s):** McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** See schematic, above

**Description:** This putative shift is similar in character to the  $\text{badw} \rightarrow \text{badu} \rightarrow \text{bidu}$  shift in Bedouin Hijazi Arabic discussed on the previous page. Indeed, the  $B \rightarrow C$  part of the shift is exactly the same, a productive process of vowel raising (/a/  $\rightarrow$  [i]) in open syllables. In this instance, the  $A \rightarrow B$  part of the shift is instantiated in a rule of epenthesis, as opposed to vocalization. Bedouin Hijazi Arabic does not tolerate final clusters, and so an epenthetic /u/ is inserted. Moreton, it appears, explicitly relates the two processes, drawing on the following comment in McCarthy (1999, p.344): “Both vocalization and epenthesis render the raising process opaque”. It should again be noted that McCarthy does not refer to either this effect, or that on the previous page, as a chain shift.

**Language:** Basaa

**ISO Code:** bas

**Language Family:** Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Northwest, A, Basaa (A.43)

**Spoken in:** Cameroon (Southern regions)

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Isomorphic

**Schematic:**  $a/\varepsilon \rightarrow e \rightarrow i$ ;  $\text{ɔ} \rightarrow o \rightarrow u$  (adapted from M)

**Environment(s):** /\_C<sub>(0)</sub>{certain suffixes}

**Source(s):** Schmidt (1996), Parkinson (1996), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *All examples from Schmidt (1996, pp.239-240)*

$A \rightarrow B$ :  $a/\varepsilon \rightarrow e$ ,  $\text{ɔ} \rightarrow o$   
[cam] → [cemha] ‘spread’  
[hæk] → [hegha] ‘create’  
[yɔŋ] → [yoŋha] ‘take’

$B \rightarrow C$ :  $e \rightarrow i$ ,  $o \rightarrow u$   
[teŋ] → [tiŋil] ‘tie’  
[tóp] → [túbùl] ‘sing’

**Description:** There are many suffixes that motivate this vowel raising pattern in Basaa, including “the passive, direct causative, simultaneous, reversive and stative extensions” (Schmidt 1996, p.240). The shifts are not assimilatory in nature. Schmidt gives an analysis of the Basaa data using linear rules. Ironically, despite listing the effect as a chain shift in the title of her article, Schmidt’s analysis does not actually treat the effect as a chain shift, as surface instances of [i, u], are derived from underlying /ɪ, ʊ/, which never surface. However, the effect in Basaa is amenable to the kind of genuine chain shift analyses that have been used to analyse metaphony processes in Romance dialects.

**Language:** Basque (Etxarri Navarrese dialect)

**ISO Code:** eus

**Language Family:** Isolate

**Spoken in:** Spain (Navarra region)

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising/Diphthongization

**Domain:** Isomorphic

**Schematic:**  $e/o \rightarrow i/u \rightarrow i^y/u^w$  (M)

**Environment(s):** /\_\_ {e}

**Source(s):** Kenstowicz & Kisseberth (1979), Hualde (1989) Kirchner (1995), Moreton & Smolensky (2002), Kawahara (2002), Moreton (2004a)

**Examples:** *All examples from Kirchner (1995, p.5)*

$A \rightarrow B$ :  $e/o \rightarrow i/u$

[seme bat] ‘son, definite’ vs. [semie] ‘son, indefinite’

[asto bat] ‘donkey, definite’ vs. [astue] ‘donkey, indefinite’

$B \rightarrow C$ :  $i/u \rightarrow i^y/u^w$

[ari bet] ‘thread, definite’ vs. [ariye] ‘thread, indefinite’

[iku bet] ‘fig, definite’ vs. [ikuwe] ‘fig’ ‘indefinite’

**Description:** Though this is not explained in Kirchner (1995), it appears that {-e} is the indefinite suffix in Etxarri Navarrese Basque. This suffix conditions raising in mid-vowels, and diphthongization in high vowels. Kirchner states that “the high vowel in [i<sup>y</sup>] and [u<sup>w</sup>] will inevitably be somewhat higher than plain [i] and [u], due to coarticulation with the off-glide” (1995, p. 5). Kirchner models the process using a constraint which he labels HIATUS RAISING, defined thus: “In V<sub>1</sub>V<sub>2</sub>, raise V<sub>1</sub> one step from its underlying height value” (p.6). This constraint interacts with a Locally-Conjoined constraint in Kirchner’s final analysis. Moreton & Smolensky (2002) present a similar Local Conjunction analysis.

**Language:** Bengali (Standard Colloquial dialect)

**ISO Code:** ben

**Language Family:** Indo-European, Indo-Iranian, Indo-Aryan, Eastern Zone, Bengali-Assamese

**Spoken in:** Bangladesh (throughout), India (Bengal region), Nepal, Singapore

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Isomorphic

**Schematic:**  $\varepsilon \rightarrow e \rightarrow i, \text{ɔ} \rightarrow o \rightarrow u$

**Environment(s):** /\_\_ {certain tense suffixes}

**Source(s):** Dasgupta (1982), Lahiri (2003), Ghosh (2001), Mahanta (2007; 2012), Nagle (2008; 2013; 2016)

**Examples:** *All examples from Mahanta (2007, p.156)*

$A \rightarrow B$ :  $\varepsilon \rightarrow e, \text{ɔ} \rightarrow o$   
dek<sup>h</sup>a ‘to see’ (nominal)  $\rightarrow$  dek<sup>h</sup>i ‘I see’  
kɔra ‘to do’ (nominal)  $\rightarrow$  kori ‘I do’

$B \rightarrow C$ :  $e \rightarrow i, o \rightarrow u$   
fek<sup>h</sup>a ‘to learn’ (nominal)  $\rightarrow$  fik<sup>h</sup>i ‘I learn’  
k<sup>h</sup>ola ‘to open’  $\rightarrow$  k<sup>h</sup>uli ‘I open’

**Description:** In verbal paradigms in Bengali, there is an alternation between the lax mid vowels and tense mid vowels, and a further alternation between the tense mid vowels and high vowels. There is some debate in the literature as to whether this constitutes a raising effect (lax (low mid)  $\rightarrow$  tense (high mid)  $\rightarrow$  high) or the reverse lowering effect. Mahanta (2007) marshals convincing arguments in favour of the former proposal. She also explicitly labels the effect in Bengali a chain shift, and uses Local Conjunction to model the shift, conjoining IDENT[ATR] with IDENT[high]. Nagle (2008) gives an account of the same effect using Optimality Theory with Candidate Chains, and Ghosh (2001) gives a more classical OT account, based on the spreading of [+ATR].

**Language:** Catalan

**ISO Code:** cat

**Language Family:** Indo-European, Italic, Romance, Italo-Western, Gallo-Iberian, Ibero-Romance, East-Iberian

**Spoken in:** Catalonia (throughout), parts of Southern France, Andorra

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Consonant Deletion

**Domain:** Adjacent

**Schematic:**  $nt\# \rightarrow n\# \rightarrow \#$  (M)

**Environment(s):** /\_\_#

**Source(s):** Mascaró (1976), Boersma (1999), Moreton & Smolensky (2002), Padgett (2002), Moreton (2004a), Neasom (2011)

**Examples:**

$A \rightarrow B$ :  $nt\# \rightarrow n\#$

/bint/  $\rightarrow$  [bin] 'twenty' (Boersma, 1999, p. 11)

/punt/  $\rightarrow$  [pun] 'point' (Mascaró, 1976, p. 86)

$B \rightarrow C$ :  $n\# \rightarrow \#$

/bin/  $\rightarrow$  [bi] 'wine' (Boersma, 1999, p. 11)

/plan/  $\rightarrow$  [pla] 'even' (Mascaró, 1976, p. 86)

**Description:** A widely cited interaction of two separate deletion rules, Mascaró (1976) comments in detail on these processes but does not explicitly relate them. Both Boersma (1999) and Moreton & Smolensky (2002) use the Catalan data as test cases for particular versions of Optimality Theory, both of which involve the local conjunction of constraints. It should, however, be noted, that both of these accounts use just one set of examples, as above. For more discussion of these and related points, see Neasom (2011; 2013), and section 5.3.1 of this thesis.

**Language:** Chemehuevi

**ISO Code:** ute

**Language Family:** Uto-Aztecan, Northern Uto-Aztecan, Numic, Southern

**Spoken in:** USA (Lower Colorado river, California)

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Deletion  
 $B \rightarrow C$ : Vowel Deletion

**Domain:** Adjacent

**Schematic:**  $V_1V_2\# \rightarrow V_1\# \rightarrow \#$  (M)

**Environment:** /\_\_#

**Source(s):** Press (1979), Moreton & Smolensky (2002), Major (2005), Moreton (2004a), Wolf (2011), Neasom (2011)

**Examples:** *all examples from Press (1979, p.26)*

$A \rightarrow B$ :  $V_1V_2\# \rightarrow V_1\#$   
/moa/  $\rightarrow$  [mo] ‘father’  
/nukwivaa/  $\rightarrow$  [nukwiva] ‘will run’

$B \rightarrow C$ :  $V_1\# \rightarrow \#$   
/paci/  $\rightarrow$  [pac] ‘daughter’

**Description:** The pattern above is well represented in Press’ extensive corpus (1979), with underlying word-final long vowels shortening and short vowels syncopating completely. However, it is important to note that these vowels never surface in Chemehuevi. Press’ rule-based account postulates that the final vowels are made voiceless and then deleted (p.20), meaning that they are never able to surface, however she is clear that there is no direct evidence for this proposition (p.13). A further complication arises in that Press also postulates a lengthening rule for word-final short vowels in monosyllables, meaning, for example, that the surface form for /moa/ should actually be [mo:].

**Language:** Chukchi

**ISO Code:** ckt

**Language Family:** Chukotko-Kamchatkan, Chukotko

**Spoken in:** Russia (Chukchi peninsula, Siberia)

**Process(es) Involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Vowel Harmony

**Domain:** Non-overlapping

**Schematic:**  $\#CCV \rightarrow \#\text{ə}CCV \rightarrow \#\text{ə}CC\text{ə}$  [Note (from Moreton 2004a): Underlying schwa can trigger harmony] (M)

**Source(s):** Bogoraz (1922), Spencer (1999), Moreton & Smolensky (2002), Moreton (2004a)

**Environment(s):**  $A \rightarrow B$ :  $/C+ \_ \_ CC$ ,  $B \rightarrow C$ : Unlimited spread within prosodic word

**Examples:** *all examples from Spencer (1999, online)*

$A \rightarrow B$ :  $CCV \rightarrow \#\text{ə}CCV$   
 $/\text{wejem} + \text{lq} + \text{n}/ \rightarrow [\text{wejem}\text{ə}\text{lq}\text{n}]$  'teeming with rivers' (section 2.2)

$B \rightarrow C$ :  $\#\text{ə}CCV \rightarrow \#\text{ə}CC\text{ə}$   
 $/\text{tələk} + \text{gərgən}/ \rightarrow [\text{təlagərgən}]$  'step, path' (section 4.1)

**Description:** In Chukchi, epenthetic schwa is used to break up consonant clusters. Syllabification usually proceeds from left-to-right, but this pattern can be broken up by morpheme boundaries (Spencer, 1999). This means that schwa will be inserted morpheme-initially, to break up CCCV clusters ( $C\text{ə}CCV$ ). In Moreton's schematic, I assume that # represents morpheme, rather than word boundaries, as Spencer's data does not include any word-initial schwa insertion. Spencer (1999) divides Chukchi vowels into two classes, labelled 'dominant' and 'recessive'. Dominant vowels trigger harmony, in which every vowel in the underlying form becomes a member of the dominant set. Schwa can be underlyingly dominant and recessive. Spencer points out that “[a]ll instances of epenthetic schwa are recessive” (1999, ch2). This means that, unlike underlying schwa, epenthetic schwa cannot trigger harmony.

**Language:** Danish

**ISO Code:** dan

**Language Family:** Indo-European, Germanic, North, East Scandinavian, Danish-Swedish, Danish-Riksmal, Danish

**Spoken in:** Denmark, Germany (Schleswig-Holstein region), Greenland

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Deaspiration/Flapping)  
 $B \rightarrow C$ : Consonant Lenition (Vocalization)

**Domain:** Isomorphic

**Schematic:**  $p^h t^s k^h \rightarrow p t/r k \rightarrow p/w \delta y/w$  (Note: Coronal  $\delta$  is an approximant in Danish)

**Environment(s):** /\_\_#

**Source(s):** Harris (2004), Hart (2010), Krämer (2012)

**Examples:** *all examples from Hart (2010, pp.21-23)*

$A \rightarrow B$ :  $p^h t^s k^h \rightarrow p t/r k$   
mikroskopi [mik<sup>h</sup>ʁosko<sup>h</sup>p<sup>h</sup>i] ‘microscopy’ vs. mikroskop [mik<sup>h</sup>ʁosko:<sup>?</sup>p]  
‘microscope’  
demokrati [demo<sup>h</sup>kʁa<sup>h</sup>t<sup>h</sup>i:<sup>?</sup>] ‘democracy’ vs. demokrat [demo<sup>h</sup>kʁa:<sup>?</sup>t]  
‘democrat’  
lakere [la<sup>h</sup>k<sup>h</sup>e:<sup>?</sup>ʁ] ‘lacquer’ (v.) vs. lak [lak] ‘lacquer’ (n.)

$B \rightarrow C$ :  $p t/r k \rightarrow p/w \delta y/w$   
hydrofobi [hy.tʁo.fo.<sup>?</sup>pi:<sup>?</sup>] ‘hydrophobia’ vs. købe [k<sup>h</sup>ø:pə] / [k<sup>h</sup>ø:w] ‘to buy’  
abbedisse [a.pe.<sup>?</sup>ti.sə] ‘abbess’ vs. abed [apeð] ‘abbot’  
køgt [k<sup>h</sup>ʁkt] → koge [k<sup>h</sup>ɔ:w] ‘to cook’

**Description:** Danish is a language in which consonant lenition is particularly pervasive. Two major treatments of Danish lenition (Harris (2004) and Hart (2010)) are clear that the kinds of contrasts that can be observed are constrained by positional effects. Both treatments illustrate that there is a foot-initial contrast between aspirated and plain plosives, and a contrast between plain plosives and approximants syllable finally. Hart, who explicitly labels the process a chain shift (similarly to Krämer 2012), uses Local Conjunction of constraints to model this effect. The motivation behind this is made clear by Hart, who asserts that constraints are conjoined to “prevent underlying forms from leniting too far” (p.34).

**Language:** Dutch (Hellendoorn dialect)

**ISO Code:** nld

**Language Family:** Indo-European, Germanic, West, Low Saxon-Low Franconian, Low Franconian

**Spoken in:** Netherlands (Hellendoorn municipality, Overijssel province)

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Place Assimilation

**Domain:** Overlapping

**Schematic:**  $ktn \rightarrow kn \rightarrow k\eta$  (M. *A more accurate version might be  $Ttn \rightarrow Tn \rightarrow TN$ , where T stands for a generic plosive*)

**Environment(s):**  $A \rightarrow B$ : /k\_\_n,  $B \rightarrow C$ : /k\_\_

**Source(s):** van Oostendorp (2004), Moreton (2004a)

**Examples:** *all examples from van Oostendorp (2004, p.2-3)*

$A \rightarrow B$ :  $ktn \rightarrow kn$   
werkten  $\rightarrow$  [wɛrkɛn] ‘(we) worked’  
hoopten  $\rightarrow$  [hoptɛn] ‘(we) hoped’

$B \rightarrow C$ :  $kn \rightarrow k\eta$   
werken  $\rightarrow$  [wɛrkɛ] ‘to work’  
hopen  $\rightarrow$  [hoptɛ] ‘to hope’

*Note: Italicized forms are **orthographic** representations from Standard Dutch, following van Oostendorp (2004, p.2-3). I assume that they match the underlying forms in Standard Dutch, although a native Western Dutch speaker (Mirjam de Jonge, pers comm.) informs me that her pronunciations are [vɛɪktə], [hoptə], [vɛɪkə], [hoptə] respectively.*

**Description:** In Hellendoorn Dutch, the /t/ that appears to be present in the Standard Dutch forms is deleted, though its presence in the underlying form is enough to block the typical process of nasal place assimilation that takes place in forms with no underlying /t/. In fact, van Oostendorp (2004, p.3) asserts that the standard assumption is that assimilation *does* take place, and a generic nasal /N/ assimilates to the coronal /t/ before it is deleted. Van Oostendorp, however, favours an account in which a coronal residue is left behind after deletion, attaching itself to the nasal.

**Language:** English (African American Detroit dialect)

**ISO Code:** eng

**Language Family:** Indo-European, Germanic, West, English

**Spoken in:** USA (Detroit, Michigan)

**Process(es) Involved:**  $A \rightarrow B$ : Monophthongization  
 $B \rightarrow C$ : Devoicing

**Domain:** Adjacent

**Schematic:**  $ard \rightarrow a^e d \rightarrow art$  (adapted slightly from M)

**Environment(s):**  $A \rightarrow B$ : /\_\_t,  $B \rightarrow C$ : /\_\_d

**Source(s):** B. Anderson (2002), Moreton (2004a)

**Examples:** None given

**Description:** No examples are given in the text of Anderson (2002). However, the article concerns the monophthongisation of the /ai/ diphthong in African American speakers in Detroit in words like, for example *tight* and *tied* (2002, p. 88). The study describes monophthongisation as common before voiced obstruents, and increasingly common before voiceless obstruents. Anderson conducts a multi-generational study, showing that in older speakers, there is no monophthongization before voiceless consonants, but within two generations monophthongization in this context has become the norm. It is not easy to see how any notion of synchronic chain shift can be adduced from the data in Anderson's article (see Chapter 5.2.1.2 for more on this).

**Language:** English

**ISO Code:** eng

**Language Family:** Indo-European, Germanic, West, English

**Spoken in:** USA, UK, Canada, Australia, South Africa (for reasons of space, only top 5 countries by population are listed)

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Deaspiration)  
 $B \rightarrow C$ : Consonant Lenition (Voicing)

**Domain:** Isomorphic

**Schematic:**  $p^h t^h k^h \rightarrow p t k \rightarrow b d g$  (M)

**Environment(s):** None suggested in Moreton's schematic

**Source(s):** Moreton (2004a)

**Examples:** None given

**Description:** No environment for this shift is given, and the process certainly is not unconditioned. In English, aspirated and plain stops appear in complementary distribution. The only time when an underlyingly aspirated stop might conceivably surface as a plain voiceless stop would be after affixation, but it appears that the reverse actually holds. Consider the contrast between British English *atom* (æ.təm) and *atomic* ə.(t<sup>h</sup>ɔm.ɪk). There is allophonic variation between [t<sup>h</sup>] and [t], but the underlying form would be standardly assumed to be /t/ rather than /t<sup>h</sup>/ in this case (see e.g. Hayes 2009, p.90) It is therefore difficult to see where the argument that plain plosives may be derived from aspirated plosives comes from.

**Language:** English (Standard Southern British dialect)

**ISO Code:** eng

**Language Family:** Indo-European, Germanic, West, English

**Spoken in:** UK (Southern England)

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Epenthesis

**Domain:** Overlapping

**Schematic:** nis  $\rightarrow$  ns  $\rightarrow$  nts (M. in his corpus, Moreton marks this as ‘very doubtful’)

**Environment(s):** /n\_\_s

**Source(s):** Donegan & Stampe (1979), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Donegan & Stampe (1979, p.154)*

$A \rightarrow B$ : nis  $\rightarrow$  ns  
/smɪstr/  $\rightarrow$  [smɪstr] ‘sinister’

$B \rightarrow C$ : ns  $\rightarrow$  nts  
/spɪnstr/  $\rightarrow$  [spɪntstr] ‘spinster’

**Description:** Donegan & Stampe (1979), in their discussion of Natural Phonology, discuss a general process of stop insertion in English that takes place between nasals and fricatives. This rule is optional, but can occur widely in the productions of some English speakers. Another rule of English, called casual speech syncope by Donegan & Stampe (1979, p.154), creates nasal/fricative clusters by deleting the vowel that occurs between them. The nasal/fricative clusters that are created by casual speech syncope do not feed stop insertion, therefore \*[sɪntstr] is not a possible production in English, even by a speaker for whom stop insertion is productive. Donegan & Stampe refer to this process as counterfeeding. However, as M&S (2002) note, this would invalidate the typological predictions of their Local Conjunction model, and they point to experimental evidence (Manuel, et al., 1992) suggesting that schwa is not actually fully deleted in the environments suggested by Donegan & Stampe.

**Language:** Finnish 1

**ISO Code:** fin

**Language Family:** Uralic, Finnic

**Spoken in:** Finland, Russia, Sweden

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Shortening  
 $B \rightarrow C$ : Lowering/Rounding/Deletion

**Domain:** Isomorphic/Adjacent

**Schematic:**  $V: \rightarrow V \rightarrow [-\text{high}]/[+\text{round}]/\emptyset$

**Environment(s):** /\_\_i

**Source(s):** Anttila (2000), Harrikarri (2000), Łubowicz (2003a; 2012)

**Examples:** *all examples from Łubowicz (2003a, p.94)*

$A \rightarrow B$ :  $V: \rightarrow V$

<i>kallii-</i>	'expensive' (sg.)	<i>kalli-i-na</i>	'expensive' (pl.)
<i>essee</i>	'essay' (sg.)	<i>esse-i-nä</i>	'essays'
<i>jää</i>	'ice'	<i>jä-i-na</i>	'ice' (pl.)
<i>tehtaa</i>	'factory'	<i>tehta-i-na</i>	'factories'

$B \rightarrow C$ :  $V \rightarrow [-\text{high}]/[+\text{round}]/\emptyset$

<i>lasi</i>	'glass'	<i>lase-i-na</i>	'glasses'	<i>kesä</i>	'summer'	<i>kes-i-nä</i>	'summers'
<i>lapse</i>	'child'	<i>laps-i-na</i>	'children'	<i>matka</i>	'trip'	<i>matko-i-na</i>	'trips'
<i>tekijä</i>	'author'	<i>tekijö-i-na</i>	'authors'	<i>asema</i>	'station'	<i>asem-i-na</i>	'stations'

**Description:** Łubowicz provides a detailed discussion of a set of related processes in Finnish. Unrounded long vowels shorten in the plural without exception, as in the first set of examples above. Underlying short vowels display a range of behaviours depending on their height. High short vowels obligatorily lower, mid short vowels obligatorily delete, whilst low short vowels either round or delete depending on more complex subregularities. Given that vowel shortening is a general process affecting all of the vowels in the series, the second part of the shift may be a reflex of the first, despite the variance in surface realizations.

**Language:** Finnish 2

**ISO Code:** fin

**Language Family:** Uralic, Finnic

**Spoken in:** Finland, Russia, Sweden

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Degemination)  
 $B \rightarrow C$ : Consonant Lenition (Spirantization/Voicing/Deletion)

**Domain:** Isomorphic

**Schematic:** pp tt kk  $\rightarrow$  p t k  $\rightarrow$  v d ø (M)

**Environment(s):** /V\_\_V

**Source(s):** Jensen & Stong-Jensen (1976), Pöchtrager (2001), Moreton & Smolensky (2002), Kiparsky (2003), Moreton (2004a), Karlsson (2008)

**Examples:** *all examples from Karlsson (2008, p.44)*

$A \rightarrow B$ : pp tt kk  $\rightarrow$  p t k

STEM	BASIC FORM	MEANING
saappaa-	saapas	'boot'
rattaa-	ratas	'wheel'
rakkaa-	rakas	'dear'

$B \rightarrow C$ : p t k  $\rightarrow$  v t ø

varpaa-	varvas	'toe'
hitaa-	hidas	'slow'
kokee-	koe	'experiment'

**Description:** The basic generalization in Finnish consonant gradation is that in sonorant contexts (intervocally or between a sonorant consonant and a vowel) consonants lenite if the following syllable is closed (for a much fuller discussion, see Karlsson (2008 pp.38-52)). This process is rarely, if ever considered a chain shift (it is not discussed as such by Jensen & Stong-Jensen (1976), Pöchtrager (2001), Kiparsky (2003) or Karlsson (2008)), however it does appear that a case could be made for a unified analysis of the phenomenon, which has been traditionally described in terms of a counterfeeding rule order, wherein the degemination rule affecting forms such as *rattaa-* ~ *ratas* necessarily occurs after the various rules affecting the series of plain stops /p t k/.

**Language:** Gbanu

**ISO Code:** gbv

**Language Family:** Niger-Congo, Atlantic-Congo, Volta-Congo, North, Adamawa-Ubangi, Ubangi, Gbaya-Manza-Ngbaka, Central

**Spoken in:** Various regions, Central African Republic

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Isomorphic

**Schematic:**  $\varepsilon \rightarrow e \rightarrow i, \text{ɔ} \rightarrow o \rightarrow u$

**Environment(s):** /\_\_ {perfective}

**Source(s):** Bradshaw (1996), Parkinson (1996)

**Examples:** *all examples from Parkinson (1996, p.69)*

$A \rightarrow B$ :  $\varepsilon \rightarrow e, \text{ɔ} \rightarrow o$

zεke	‘sift’ (imperf.)	zeke	‘has sifted’ (perf.)
gɔmɔ	‘chop’ (imperf.)	gomɔ	‘has chopped’ (perf.)

$B \rightarrow C$ :  $e \rightarrow i, o \rightarrow u$

zele	‘hear’ (imperf.)	zile	‘has heard’ (perf.)
tombo	‘send’ (imperf.)	tumbo	‘has sent’ (perf.)

**Description:** Parkinson (1996) and Bradshaw (1996) argue that the perfective form in Gbanu is created by taking the imperfective and raising the preceding non-low vowel by one step. Parkinson represents height via the addition of [closed] features (in the style of Clements (1991)). It is Parkinson’s contention that the morpheme that changes the imperfective into a perfective form is a [closed] feature that is added to the first vowel of the word (V1). Interestingly, for monosyllabic imperfective forms such as [pe], the perfective form is [peɛ] as opposed to [pe] (example from Parkinson (1996, p. 69)). Bradshaw (1996) treats the second vowel as an extra mora that is added as part of the perfective morphology.

**Language:** Hidatsa

**ISO Code:** hid

**Language Family:** Siouan-Catawban, Siouan, Missouri River Siouan

**Spoken in:** Fort Berthold Reservation, North Dakota, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Deletion  
 $B \rightarrow C$ : Vowel Deletion

**Domain:** Adjacent

**Schematic:**  $V_1V_2\# \rightarrow V_1\# \rightarrow \#$  (M)

**Source(s):** Harris (1942), Moreton & Smolensky (2002), Boyle (2002), Moreton (2004a), Wolf (2011), Neasom (2011)

**Environment(s):** /\_\_#

**Examples:** *all examples from Harris (1942, p.171)*

$A \rightarrow B$ :  $V_1V_2\# \rightarrow V_1\#$   
/kikua/ ‘set a trap’ (infin.)  $\rightarrow$  [kiku] ‘set a trap!’ (imper.)  
/ika:/ ‘look’ (infin.)  $\rightarrow$  [ika] ‘look!’ (imper.)

$B \rightarrow C$ :  $V_1\# \rightarrow \#$   
/cixi/ ‘jump’ (infin.)  $\rightarrow$  [cix] ‘jump!’ (imper.)

**Description:** In Hidatsa, the imperative of a verb is formed by the word-final syncope of a mora. This analysis is first proposed in Harris (1942), but an opposing, chain shift analysis is formulated in Moreton & Smolensky (2002), in which the effect is treated with Local Conjunction. Word-final consonant deletion is obligatory, but a local conjunction of the faithfulness constraint MAX stops two word-final consonants being deleted, or a short vowel being deleted entirely. For more extensive discussion, see Chapter 5.3.4.

**Language:** Hmong (Dananshan dialect)

**ISO Code:** cqd

**Language Family:** Hmong-Mien, Hmongic, Chuangqiandian

**Spoken in:** Dananshan village, Guizhou province, China

**Process(es) Involved:**  $A \rightarrow B$ : Tone Sandhi  
 $B \rightarrow C$ : Tone Sandhi

**Domain:** Isomorphic

**Schematic:**  $\uparrow H \rightarrow H \rightarrow M$

**Environment(s):** /{HM,ML}\_\_

**Source(s):** F. Wang (1985), Niederer (1998), Mortensen (2006)

**Examples:** *all examples from Mortensen (2006, p.73)*

$A \rightarrow B$ :  $\uparrow H \rightarrow H$

ntou <sup>HM</sup> ‘cloth’	+	sa <sup>↑H</sup> ‘blue’	→	ntou <sup>HM</sup> sa <sup>H</sup> ‘blue cloth’
ku <sup>ML</sup> ‘trench’	+	tse <sup>↑H</sup> ‘house’	→	ku <sup>ML</sup> tse <sup>H</sup> ‘sewer’

$B \rightarrow C$ :  $H \rightarrow M$

au <sup>HM</sup> ‘two’	+	pua <sup>H</sup> ‘hundred’	→	au <sup>HM</sup> pua <sup>M</sup> ‘two hundred’
nplon <sup>ML</sup>	+	nton <sup>H</sup> ‘tree’	→	nplon <sup>ML</sup> nton <sup>M</sup> ‘tree leaf’

**Description:** In the Danashan dialect of Hmong, there is a tone sandhi process conditioned by a preceding falling tone (HM or ML). This tone causes the following tone to be represented one step down from its citation form. Super high tones ( $\uparrow H$ ) become high (H) in the sandhi context, whilst high tones become mid (M). Mortensen (2006) suggests that the motivation for this putative shift was initially a phonetic, co-articulatory effect created by the proximity of the falling tone, which over time grew more pronounced and at some stage became phonologized.

**Language:** Hmong (Shuijingping dialect) 1

**ISO Code:** hmn

**Language Family:** Hmong-Mien, Hmongic

**Spoken in:** South Guizhou Province, China

**Process(es) Involved:**  $A \rightarrow B$ : Tone Sandhi  
 $B \rightarrow C$ : Tone Sandhi

**Domain:** Isomorphic

**Schematic:** HM  $\rightarrow$   $\uparrow$ H  $\rightarrow$  H

**Environment(s):** /{M,L}\_\_

**Source(s):** Mortensen (2006)

**Examples:** *all examples from Mortensen (2006, p.85)*

$A \rightarrow B$ : HM  $\rightarrow$   $\uparrow$ H

ka <sup>M</sup> ‘medicine’	+	tcə <sup>HM</sup> ‘liquor’	$\rightarrow$	ka <sup>M</sup> tcə <sup><math>\uparrow</math>H</sup> ‘brewer’s yeast’
na <sup>L</sup> ‘bag’	+	ntə <sup>HM</sup> ‘book’	$\rightarrow$	na <sup>L</sup> ntə <sup><math>\uparrow</math>H</sup> ‘book bag’

$B \rightarrow C$ :  $\uparrow$ H  $\rightarrow$  H

he <sup>M</sup> ‘chicken’	+	he <sup><math>\uparrow</math>H</sup> ‘crow’	$\rightarrow$	he <sup>M</sup> he <sup>H</sup> ‘cock’s crow’
ti <sup>L</sup> ‘RECIP’	+	he <sup><math>\uparrow</math>H</sup> ‘curse’	$\rightarrow$	ti <sup>L</sup> he <sup>H</sup> ‘quarrel’

**Description:** In the Shuijingping dialect of Hmong, high-mid tones (HM) become super-high ( $\uparrow$ H) when following mid (M) or low (L) tones. Simultaneously, tones which are super-high in citation forms lower to become high (H) in the same context. There is no specific phonological context in which this shift occurs. The triggers are the same, however instead of raising, lowering is observed. Indeed, the lowering brings the super-high tone back in the direction of the HM tone that forms the first element of the shift. Mortensen claims that this shift is based on a historically earlier shift that took the form  $\uparrow$ H  $\rightarrow$  H  $\rightarrow$  HM. He suggests that a diachronic, circular chain shift has occurred (2006, p.85).

**Language:** Hmong (Shuijingping dialect) 2

**ISO Code:** hmh

**Language Family:** Hmong-Mien, Hmongic

**Spoken in:** South Guizhou province, China

**Process(es) Involved:**  $A \rightarrow B$ : Tone Sandhi  
 $B \rightarrow C$ : Tone Sandhi

**Domain:** Isomorphic

**Schematic:** LML  $\rightarrow$  MH  $\rightarrow$  ML (Note: There is what appears to be a transcription error in the data. This is the schematic that Mortensen gives, however in all examples ML is realized as LM. In terms of Mortensen's discussion, a realization of ML makes more sense, so I have adapted the examples accordingly)

**Environment(s):** /{M,L}\_\_

**Source(s):** Mortensen (2006)

**Examples:** all examples from Mortensen (2006, p.86)

$A \rightarrow B$ : LML  $\rightarrow$  MH

kua <sup>M</sup> 'bug'	+	ntsɔ <sup>LML</sup> 'louse'	$\rightarrow$	kua <sup>M</sup> ntsɔ <sup>MH</sup> 'head louse'
mu <sup>L</sup> 'NEG'	+	laŋ <sup>LML</sup> 'grow'	$\rightarrow$	mu <sup>L</sup> laŋ <sup>MH</sup> 'not grow'

$B \rightarrow C$ : MH  $\rightarrow$  ML

ʔei <sup>M</sup> 'one'	+	li <sup>MH</sup> 'month'	$\rightarrow$	ʔei <sup>M</sup> li <sup>ML</sup> 'one month'
mu <sup>L</sup> 'NEG'	+	naŋ <sup>MH</sup> 'believe'	$\rightarrow$	mu <sup>L</sup> naŋ <sup>ML</sup> 'not believe'

**Description:** In the Shuijingping dialect of Hmong, low-mid-low (LML) citation tones become mid-high (MH) tones when directly preceded by mid (M) or low (L) tones. Simultaneously, citation MH tones become mid-low (ML) tones when directly preceded by the same trigger tones, M and L. This shift thus involves both raising and lowering. It also involves the loss of a melody tone (LML  $\rightarrow$  MH), making it unique amongst the Hmongic tone shifts discussed by Mortensen. There is no obvious phonological motivation for these shifts, which leads Mortensen to postulate a more abstract explanation. Mortensen suggests that an earlier, more phonetically natural shift (similar to that which exists synchronically in Danashan Hmong) existed in Shuijingping Hmong, but over time became less natural.

**Language:** Hmong (Xinzhai dialect)

**ISO Code:** hmn

**Language Family:** Hmong-Mien, Hmongic

**Spoken in:** South Guizhou province, Guizhou

**Process(es) Involved:**  $A \rightarrow B$ : Tone Sandhi  
 $B \rightarrow C$ : Tone Sandhi

**Domain:** Isomorphic

**Schematic:** MH  $\rightarrow$  LM  $\rightarrow$  ML

**Environment(s):** /{H, M, ML}\_\_

**Source(s):** Mortensen (2006)

**Examples:** *all examples from Mortensen (2006, p.84)*

$A \rightarrow B$ : MH  $\rightarrow$  LM

zei <sup>M</sup> ‘honey’	+	mu <sup>MH</sup> ‘bee’	$\rightarrow$	zei <sup>M</sup> mu <sup>LM</sup> ‘honey bee’
nou <sup>LM</sup> ‘day’	+	no <sup>MH</sup> ‘this’	$\rightarrow$	nou <sup>LM</sup> no <sup>LM</sup> ‘today’

$B \rightarrow C$ : LM  $\rightarrow$  ML

tou <sup>M</sup> ‘boy’	+	zou <sup>LM</sup> ‘young’	$\rightarrow$	tou <sup>M</sup> zou <sup>ML</sup> ‘young boy’
soŋ <sup>H</sup> ‘fat’	+	mpo <sup>LM</sup> ‘pig’	$\rightarrow$	soŋ <sup>H</sup> mpo <sup>ML</sup> ‘lard’

**Description:** In the Xinzhai dialect of Hmong, mid-high contour tones (MH) lower to become low-mid (LM) after high (H), mid (M) and low-mid tones. In the same context, citation LM tones raise to become ML. This shift cannot be viewed as phonetically natural in any sense for two reasons. The first is that the triggers for the shift run from H to LM, suggesting that they do not share some crucial triggering feature. The second is that the shift contains both lowering and raising elements. As is shown in other shifts based on tone and even tone circles, the general tendency in processes of this kind is lowering, with raising in circle shifts only occurring from a citation form that takes the lowest possible tone in the language.

**Language:** Icelandic

**ISO Code:** isl

**Language Family:** Indo-European, Germanic, North, West Scandinavian

**Spoken in:** Iceland

**Process(es) Involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Umlaut

**Domain:** Non-overlapping

**Schematic:** aCr#  $\rightarrow$  aCur#  $\rightarrow$  öCur# (M)

**Environment(s):**  $A \rightarrow B$ : /C\_\_r,  $B \rightarrow C$ : /\_\_C<sub>(0)</sub>u

**Source(s):** Kenstowicz (1994), Moreton (2004a)

**Examples:** *all examples from Kenstowicz (1994, p.80)*

ACC. SG	NOM.SG	DAT. PL	GLOSS
hatt	hatt-ur	hött-um	'hat'
dal	dal-ur	döl-um	'valley'
stað	stað-ur	stöð-um	'place'

**Description:** Kenstowicz, following Orešnik (1972), suggests that the nom.sg suffix (but not the dat.pl suffix) has an epenthetic [u], meaning that its underlying representation is in fact /-r/. This is borne out by examples such as *bæ-r* 'farmhouse' (nom. sg., Kenstowicz 1994, (p.79)), where [u] does not surface. The putative shift rests on a rule of umlaut, which is conditioned by suffixes containing the high vowel /u/. These suffixes condition raising from /a/  $\rightarrow$  [ö], as in the dat.pl examples above. However, umlaut does not affect the nom.sg forms. Kenstowicz suggests that this is because the umlaut rule is ordered before the epenthesis rule (see 5.3.3 for further discussion).

**Language:** Inupiaq (Barrow dialect)

**ISO Code:** esi

**Language Family:** Eskimo-Aleut, Eskimo, Inuit-Inupiaq

**Spoken in:** Barrow, Alaska, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Fronting  
 $B \rightarrow C$ : Palatalization

**Domain:** Overlapping

**Schematic:**  $iC_{(0)}l \rightarrow iC_{(0)}l \rightarrow iC_{(0)}\lambda$  (M)

**Environment(s):**  $A \rightarrow B$ : /\_\_C<sub>(0)</sub>l,  $B \rightarrow C$ : /iC<sub>(0)</sub>\_\_

**Source(s):** Kaplan (1981), Archangeli & Pulleyblank (1994), McCarthy (2002), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *Examples from McCarthy (2002, p.210)*

$A \rightarrow B$ :  $iC_{(0)}l \rightarrow iC_{(0)}l$   
/tiŋi + vluni/  $\rightarrow$  [tiŋivluni] ‘to be able to take flight’

$B \rightarrow C$ :  $iC_{(0)}l \rightarrow iC_{(0)}\lambda$   
/niri + vluni/  $\rightarrow$  [nirivλuni] ‘to be able to eat’

**Description:** Kaplan (1981) postulates that there are two URs for surface [i] in Barrow Inupiaq, which he defines as /i/ and /i/ (reported in Archangeli & Pulleyblank (Grounded Phonology, 1994, p. 82)). To my knowledge, there is no substantive phonetic difference between these two forms. Underlying /i/ forms trigger palatalization of laterals (see the  $B \rightarrow C$  forms above), but underlying /i/ does not. McCarthy (2002) analyses the data using Optimality Theory, using Comparative Markedness constraints. Comparative Markedness has been used to model chain shifts by both McCarthy (2002) and others (e.g., Dinnsen & Farris-Trimble (2008)), but McCarthy does not refer to the effect in Barrow Inupiaq as a chain shift.

**Language:** Irish 1

**ISO Code:** gle

**Language Family:** Indo-European, Celtic, Insular, Goidelic

**Spoken in:** Ireland

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Voicing)  
 $B \rightarrow C$ : Nasalization

**Domain:** Isomorphic

**Schematic:**  $p/t/k \rightarrow b/d/g \rightarrow m/n/\eta$

**Environment(s):** Various morphosyntactic environments (see below)

**Source(s):** Ní Chiosain (1991), Rice (1993), Grijzenhout (1995), Gnanadesikan (1997), Trommer (2009)

**Examples:** *from Ní Chiosain 1991, reproduced in Gnanadesikan (1997, p.97)*

$A \rightarrow B$ :  $t \rightarrow d$

/t'ax/ 'a house'  $\rightarrow$  [ə d'ax] 'their house'

$B \rightarrow C$ :  $d \rightarrow n$

/dorəs/ 'a door'  $\rightarrow$  [ə norəs] 'their door'

/dorəs/ 'a door'  $\rightarrow$  [s'axt norəs] 'seven doors'

**Description:** Gnanadesikan explicitly labels this pattern, traditionally termed eclipsis, as a chain shift (1997, p.96). Word-initial underlying voiceless plosives voice, whilst underlying voiced plosives become nasalized in a wide variety of morphosyntactic contexts. The triggers are all morphological, but the effects are phonologically regular. Gnanadesikan's OT analysis assumes that the motivation for eclipsis is morphological; a fully-faithful output violates her highly-ranked constraint MORPH-REAL, which assigns a violation to any candidate that fails to undergo eclipsis in the contexts described above (see 5.4.1 for more discussion).

**Language:** Irish 2

**ISO Code:** gle

**Language Family:** Indo-European, Celtic, Insular, Goidelic

**Spoken in:** Ireland

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Spirantization)  
 $B \rightarrow C$ : Consonant Lenition (Debuccalization/Deletion)

**Domain:** Isomorphic

**Schematic:**  $ptk \rightarrow fhx \rightarrow \emptyset h?$  (RotB<sup>20</sup> for  $x \rightarrow ?$ ) (M)

**Environment(s):** Various morphosyntactic environments

**Source(s):** Ní Chiosain (1991), Rice (1993), Gnanadesikan (1997), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Ní Chiosain 1991, reproduced in Gnanadesikan (1997, p.191)*

$A \rightarrow B$ :  $p \rightarrow f$   
 $/in + po:stə/ \rightarrow [in'fo:stə]$  'marriageable'

$B \rightarrow C$ :  $f \rightarrow \emptyset$   
 $/mə fi:əcəl/ \rightarrow [mə i:əkəl]$  'my tooth'  
 $/t'ri: fiəcəl/ \rightarrow [t'ri: i:əkəl]$  'three teeth'

**Description:** This pattern is part of a larger set of lenition processes in Irish (see 5.4.1 for more discussion). Whilst many plosives and fricatives lenite, only in the labial series is there a clear chain shift (at least from the data presented in Gnanadesikan (1997, pp.190-191)). More common are complete neutralizations ( $/t/$  and  $/s/ \rightarrow [h]$ , for instance), or partial lenition trajectories which have only one step ( $/k/ \rightarrow [x]$ ). The observed lenition effects appear to be the result of a wide variety of morphological processes (Gnanadesikan (1997, p. 190)).

---

<sup>20</sup> RotB here stands for 'Richness of the Base'. For extensive discussion of this, see section 5.2.3

**Language:** Italian (Servigliano dialect)

**ISO Code:** ita

**Language Family:** Indo-European, Italic, Romance, Italo-Western, Italo-Dalmatian

**Spoken in:** Fermo province, Marche region, Italy

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Isomorphic

**Schematic:**  $\varepsilon \rightarrow e \rightarrow i, \text{ɔ} \rightarrow o \rightarrow u$

**Environment(s):** / \_\_C<sub>(0)</sub>{i/u}

**Source(s):** Kaze (1989), Parkinson (1996), Nibert (1998), Cole (1998), Mascaró (2011), Calabrese (2011)

**Examples:** *all examples from Parkinson (1996, p.39). Stress always on penult*

$A \rightarrow B$ :  $\varepsilon \rightarrow e, \text{ɔ} \rightarrow o$

modɛst-a	‘modest’ (fem. sg.)	modest-u	‘modest’ (masc. sg.)
mɔr-e	‘he dies’	mor-i	‘you die’

$B \rightarrow C$ :  $e \rightarrow i, o \rightarrow u$

kred-o	‘I believe’	krid-i	‘you believe’
fjor-e	‘flower’	fjur-i	‘flowers’

**Description:** This is an example of a metaphony process in the Southern Italian dialect of Servigliano. There are many approaches to this kind of metaphony in the autosegmental literature (see Calabrese (2011) for full references, and 5.4.1 onwards for more discussion). In general, these approaches do not refer to the Servigliano effect as a chain shift. However, Parkinson (1996) lists the process as an instance of a vowel raising chain shift. Parkinson models raising shifts via the addition of a feature [closed], based on Clements (1991) treatment of vowel height in Bantu languages. In Parkinson’s (1996) interpretation of the data, mid-low vowels have one [closed] feature, and high-mid vowels have two (see 5.4.3.2.3 for more discussion).

**Language:** Italian (Central and Southern vernacular dialects)

**ISO Code:** ita

**Language Family:** Indo-European, Italic, Romance, Italo-Western, Italo-Dalmatian

**Spoken in:** Southern and Central Italy

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Voicing)  
 $B \rightarrow C$ : Nasalization

**Domain:** Isomorphic (but note environments)

**Schematic:** T  $\rightarrow$  D  $\rightarrow$  N

**Environment(s):**  $A \rightarrow B$ : / [+cons, +son] \_\_,  $B \rightarrow C$ : / [+nas] \_\_

**Source(s):** Chapallaz (1979), Gnanadesikan (1997)

**Examples:** *all examples from Gnanadesikan (1997, pp.74-75)*

T  $\rightarrow$  D

Standard Italian	CSVD Italian	Spelling	Gloss
[ampjo]	[ambjo]	<ampio>	'wide, roomy'
[mantʃa]	[mandʒa]	<mancia>	'tip'
[miltʃa]	[mildʒa]	<milza>	'spleen'

D  $\rightarrow$  N

[gamba]	[gamma]	<gamba>	'leg'
[mondo]	[monno]	<mondo>	'world'
[lungo]	[luŋŋo]	<lungo>	'length'

**Description:** Gnanadesikan begins her chapter on chain shifting with this example, which she asserts is common in “the vernacular dialects of much of southern and central Italy” (p.74). The generalization is that, between a sonorant and a vowel, underlyingly voiceless plosives surface as their voiced counterparts, whilst underlyingly voiced plosives nasalize on the surface. This pattern, however, is not entirely symmetrical. The first stage of the shift is more general than the second. As Gnanadesikan points out, voiceless plosives voice between *any sonorant* and a vowel. The nasalization process, unsurprisingly, occurs only in the environment of a nasal.

**Language:** Japanese (Tokyo dialect)

**ISO Code:** jpn

**Language Family:** Japonic

**Spoken in:** Tokyo, Japan

**Process(es) Involved:**  $A \rightarrow B$ : Nasalization  
 $B \rightarrow C$ : Rendaku

**Domain:** Non-overlapping

**Schematic:** ...togi  $\rightarrow$  ...tonji  $\rightarrow$  ...donji (M)

**Environment(s):**  $A \rightarrow B$ : /[-son, -voi]V \_\_,  $B \rightarrow C$ : / \_\_ V ŋ

**Source(s):** Ito & Mester (1997), McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a), Labrune (2012)

**Examples:**

$A \rightarrow B$ : ...togi  $\rightarrow$  ...tonji  
/hasami-toGi/  $\rightarrow$  [hasamitonji] ‘knife-grinder’ (from McCarthy (1999, p. 14))

$B \rightarrow C$ : ...tonji  $\rightarrow$  ...donji  
None given in sources above

**Description:** Whether this effect can be considered a chain shift is very much dependent on the representations of  $g$  and  $\eta$ . McCarthy (1999) treats both surface [g] and [ŋ] as allophones of an underlying form /G/ (based on Ito & Mester (1997)). [g] surfaces word-initially and [ŋ] word-medially or word-finally. On this account, the schematic provided in Moreton is incoherent, as there is never a fully specified /g/ in the underlying form. The second part of the shift is similarly complex. As McCarthy notes, “In Japanese, voiced obstruents dissimilatorily block *rendaku* (‘sequential voicing’) but sonorants, though also voiced, do not” (McCarthy, 1999, p. 14). This is commonly referred to as Lyman’s Law. Importantly, however, instances of [ŋ] derived from what we shall, for expedience’s sake, call /G/, have the same rendaku blocking facility as other voiceless consonants. They do not pattern with other nasals in this respect, which is why a chain shift account is postulated.

**Language:** Karok

**ISO Code:** kyh

**Language Family:** Isolate

**Spoken in:** Klamath river, Northern California, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Shortening  
 $B \rightarrow C$ : Vowel Shortening

**Domain:** Isomorphic

**Schematic:**  $V:: \rightarrow V: \rightarrow V$  (M)

**Environment(s):** /\_\_# {second part of compound}

**Source(s):** Bright (1957), Moreton & Smolensky (2002), Moreton (2004a), Levi (2008)

**Examples:** *all examples from Bright (1957, p.36)*

$A \rightarrow B$ :  $V:: \rightarrow V:$

/ʔi::n/ 'falls' + /pí:t/ 'new' → [ʔi:npi:t] 'new falls (a place name)'

/pihnî::č/ 'old man' + /xusʔé:θa:n/ 'person who takes care' → [pihni:č-xusʔé:θa:n] 'nurse for an old man'

$B \rightarrow C$ :  $V: \rightarrow V$

/ʔi:npi:t/ 'sand bar' + /θú:f/ 'creek' → [ʔinpítθú:f] 'Sandy Bar Creek'

/axvâ:h/ 'head' + /x:us/ 'smooth' → [axvâh-xu:s] 'bald-headed'

**Description:** Bright (1957, p.34) discusses a rule in Karok in which /v/ and /y/ are deleted intervocalically. The fusion that results from this process creates a class of vowels which Bright labels 'double-long'. The key characteristic of these vowels is that they do not undergo another process in Karok phonology. In the first part of compounds, there is a general rule that long vowels shorten, as in the  $V: \rightarrow V$  examples above. This would seem to suggest that the above schematic and examples are something of a simplification, and that a more accurate schematic would be  $Vv/yV \rightarrow V: \rightarrow V$ . This schematic would allow for the same generalizations to be captured, but would not force the creation of a set of double-long vowels. This allows Bright's analysis to accord more easily with more recent analyses, such as Levi (Phonemic vs. derived glides, 2008, p. 1961), which only acknowledges long and short vowels in Karok.

**Language:** Kayardild

**ISO Code:** gyd

**Language Family:** Australian, Tangic

**Process(es) Involved:**  $A \rightarrow B$ : Nasalization  
 $B \rightarrow C$ : Place Assimilation

**Domain:** Isomorphic

**Schematic:**  $c \rightarrow \eta \rightarrow n$

**Environment(s):** / \_\_+m

**Source(s):** Round (2010)

**Example(s):** *all examples from Round (2010, pp.711-712)*

$A \rightarrow B$ :  $c \rightarrow \eta$   
/waŋɟic-mara/  $\rightarrow$  [waŋɟɨmara] *warnginy-marra* ‘one-UTIL’

$B \rightarrow C$ :  $\eta \rightarrow n$   
/kuɒŋ-mara/  $\rightarrow$  [kuanmara] *kuwan-marra* ‘firestick-UTIL’

**Description:** Kayardild is extremely susceptible to morphophonological alternations at the boundary between two morphemes, which Round (2010) schematises as  $m_1+m_2$ . There are three distinct classes of such alternations, which Round terms ‘regular’, ‘deleting’, and ‘leniting’, and which class of alternation will be instantiated depends on the character of  $m_2$  (Round 2010, p.144-145). Even within one class of alternations, Round notes that chain shifts occur. The example that he gives is in the most common, ‘regular’ class of alternations. Before the initial {-m} of the utilitive morpheme, underlying palatal plosives nasalize. Simultaneously, underlying palatal nasals change in character, becoming surface alveolar nasals. The  $A \rightarrow B$  part of the shift is a straightforward nasalization, as both underlying and surface forms share a place of articulation. However, the  $B \rightarrow C$  part of the shift is not so phonetically natural, as underlying /ŋ/ does not fully assimilate to following {-m}.

**Language:** Kikuria

**ISO Code:** kuj

**Language Family:** Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Central, J, Logooli-Kuria (E.43)

**Spoken in:** Tanzania, Kenya

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Isomorphic (but note environments)

**Schematic:**  $\varepsilon \rightarrow e \rightarrow i, \text{ɔ} \rightarrow o \rightarrow u$

**Environment(s):**  $A \rightarrow B$ : /\_\_C<sub>(0)</sub>i,u  $B \rightarrow C$ : /\_\_C<sub>(0)</sub>i, u, j, w

**Source(s):** Cammenga (1994), Parkinson (1996), Chacha & Odden (1998)

**Examples:** *all examples from Parkinson (1996, p.57), tones omitted for clarity*

$A \rightarrow B$ :  $\varepsilon \rightarrow e, \text{ɔ} \rightarrow o$

INFINITIVE	AGENTIVE	GLOSS
oko-rɔg-a	omo-rog-i	‘witch’
ogo-terek-a	omo-terek-i	‘brew’

$B \rightarrow C$ :  $e \rightarrow i, o \rightarrow u$

ogo-tačor-a	omo-tačur-i	‘untie’
ogo-teget-a	umu-tigit-i	‘be late’

**Description:** The Bantu language Kikuria has a complex system of vowel harmony, described in detail in Cammenga (1994) and Chacha & Odden (1998). Underlying / $\varepsilon$   $\text{ɔ}$ / are realized on the surface as [e o] in the presence of high vowels. In the second, underlying /e o/ are realized as [i u] in the presence of high vowels, glides and certain palatal consonants (see Chacha and Odden 1998 (sections 2 and 3)). This overlapping environment, which seems to be connected to the feature [+high], has led Cammenga (1994) and Parkinson (1996) to present the effects in Kikuria as being reflexes of the same process. Chacha & Odden (1998) are more equivocal, asserting that they see no conclusive evidence for unifying the processes, but also no conclusive evidence for not doing so.

**Language:** Manya

**ISO Code:** mzi

**Family:** Niger-Congo, Mande, Western, Central-Southwestern, Central, Manding-Jogo, Manding-Vai, Manding-Mokole, Manding, Manding-East, Southeastern Manding

**Spoken in:** Liberia, Guinea

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Voicing)  
 $B \rightarrow C$ : Nasalization

**Domain:** Isomorphic

**Schematic:**  $f \rightarrow v \rightarrow m$

**Environment(s):** Various morphosyntactic triggers (see below)

**Source(s):** Heydorn (1943-44) Gnanadesikan (1997)

**Examples:** *all examples from Gnanadesikan (1997, p.114)*

$A \rightarrow B$ :  $f \rightarrow v$

[fɛ] ‘direction’ vs. [vé] ‘to me’  
/b̃ó + fila/  $\rightarrow$  [bó vila] ‘to build a house’

$B \rightarrow C$ :  $v \rightarrow m$

*No examples given*

**Description:** Gnanadesikan (1997) discusses an eclipsis process in the Liberian language Manya, explicitly stating that the effect is a chain shift. Eclipsis has two separate triggers in Manya, one morphological (“it replaces the first person singular pronoun *né* in possessives, direct objects, and objects of postpositions” (Gnanadesikan 1997, p.113), and one phonological (“after a nasalized vowel within an NP or VP” (Gnanadesikan 1997, p.113). The important generalization is that voiceless obstruents become voiced in their eclipsed form, whilst underlyingly voiced obstruents nasalize. This forms a coherent trajectory of an increase along what Gnanadesikan calls the inherent voicing scale, essentially an analogue for the sonority scale.

**Language:** Misantla Totonac

**ISO Code:** tlc

**Language Family:** Totonacan, Totonac

**Spoken in:** Western Central Mexico

**Process(es) Involved:**  $A \rightarrow B$ : Debuccalization  
 $B \rightarrow C$ : Vowel Lowering

**Domain:** Adjacent

**Schematic:**  $u(:)\text{/}i(:)\text{/} \rightarrow u(:)\text{/}i(:)\text{/} \rightarrow \text{ɔ}(:)\text{/}e(:)\text{/}$  (M. *More accurately:*  $u(:)\text{/}i(:)\text{/} \rightarrow u(:)\text{/}i(:)\text{/}h \rightarrow \text{ɔ}(:)\text{/}h/e(:)\text{/}h$ )

**Environment(s):**  $/(q, h)\_\_(q, h)$

**Source(s):** MacKay (1994), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from MacKay (1994, p.387)*

$A \rightarrow B$ :  $u(:)\text{/}i(:)\text{/} \rightarrow u(:)\text{/}i(:)\text{/}h$   
 $/ku\text{č}i\text{/} \rightarrow [kú\text{č}ih]$  ‘knife’

$B \rightarrow C$ :  $u(:)\text{/}i(:)\text{/}h \rightarrow \text{ɔ}(:)\text{/}h/e(:)\text{/}h$   
 $/suqunaa/ \rightarrow [sɔxɔnáa]$  ‘pretty’  
 $/qiin-šii-na/ \rightarrow [qeiñšɛhna]$  ‘roots’

**Description:** In Misantla Totonac, there is a general rule of lowering both before and after what MacKay calls the ‘postvelar’ consonants  $/q/$  and  $/h/$ . Both front and back high vowels of whatever length are lowered as a result of this rule, which does not affect lower vowels (there is no process whereby  $/ɛ/ \rightarrow [a]$ , for example). This lowering interacts with an independent process in which a velar lateral ( $/\text{/}$ ) becomes  $[h]$  syllable finally (MacKay, 1994, p. 399), with the effect that surface instances of  $[h]$  that are derived from  $/\text{/}$  do not trigger lowering. MacKay gives the late application of the  $/\text{/} \rightarrow [h]$  rule as a reason for the failure of lowering to apply in these contexts. As MacKay’s analysis is written in an autosegmental rule-based framework, this can be represented as a classic counterfeeding process.

**Language:** Mohawk

**ISO Code:** moh

**Language Family:** Iroquoian, Northern Iroquoian, Five Nations-Huronian-Susquehannock, Five Nations-Susquehannock, Mohawk-Oneida

**Spoken in:** Various regions, Canada, Northern New York, USA

**Process(es) Involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Stress Assignment

**Domain:** Non-overlapping

**Schematic:** ...VC?#  $\rightarrow$  ...'VCe?  $\rightarrow$  ...VC'e? (M. NB: *There are no examples of a final stressed /e/, as would be suggested by the final part of the schematic, in Postal (1969), Halle & Clements (1983), or Rawlins (2006)*)

**Environment(s):**  $A \rightarrow B$ : /C\_\_?,  $B \rightarrow C$ : \_\_#

**Source(s):** Postal (1969), Halle & Clements (1983), Moreton & Smolensky (2002), Moreton (2004a), Rawlins (2006)

**Examples:**

$A \rightarrow B$ : ...VC?#  $\rightarrow$  ...'VCe?  
/wa?+hra+ket+?/  $\rightarrow$  [wəh'a:gedə?] 'he scraped' (Postal, 1969, p. 295)  
/o+wis+?/  $\rightarrow$  ['o:wize?] 'ice, glass' (Halle & Clements, 1983, p. 121)

$B \rightarrow C$ : ...'VCe?  $\rightarrow$  ...VC'e?  
/wake+nuhwe?+u+ne?/  $\rightarrow$  [wagenuhwe?'u:ne?] 'I had liked it'  
(Halle & Clements, 1983, p. 121)  
/ã+k+hãte+?/  $\rightarrow$  [ãkh'ã:de?] 'I shall go ahead'  
(Halle & Clements, 1983, p. 121)

**Description:** In Mohawk, as originally described by Postal (1969), there is an asymmetry between the predominant penultimate stress pattern, and stress in words in which an epenthetic vowel has been inserted. In such words, stress falls on the antepenult. An opaque, counterfeeding rule order is suggested in Postal, but more recent accounts (e.g., Rawlins (2006)), do not believe that opacity is present in Mohawk stress assignment. Rawlins uses an account based on syllable structure to describe the data, suggesting that "if Mohawk has moraic trochaic feet...Mohawk stress-epenthesis interaction falls into place with no unusual prosodic structures" (Rawlins, 2006, p. 2).

**Language:** Mwera

**ISO Code:** mjh

**Language Family:** Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Central, N, Tumbuka (N.201)

**Spoken in:** Tanzania

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Voicing)  
 $B \rightarrow C$ : Consonant Deletion

**Domain:** Isomorphic

**Schematic:**  $mp \rightarrow mb \rightarrow m$  (M. Note: *The rule is actually more general than this, and can be more accurately represented as  $NT \rightarrow ND \rightarrow N$* )

**Environment(s):** / [+nas]\_\_

**Source(s):** Harries (1950), Kenstowicz & Kisseberth (1977), Noske, Schinkel, and Smith (1982), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Noske, Schinkel and Smith (1982, p.402)*

$A \rightarrow B$ :  $NT \rightarrow ND$

/n+kuya/  $\rightarrow$  [ŋ-guya] ‘cape beans’ (cf. /lu + kuya/  $\rightarrow$  [lu-kuya] ‘cape bean’)

$B \rightarrow C$ :  $ND \rightarrow N$

/n+gomo/  $\rightarrow$  [ŋ-omo] ‘lips’ (cf. /lu + gomo/  $\rightarrow$  [lu-gomo] ‘lip’)

**Description:** In Mwera, there is a proscription against sequences made up of a nasal and a voiceless obstruent. The repair strategy employed by Mwera is to voice the obstruent. Explanations based on co-articulatory (e.g., Halpert (2012)) or perceptual (e.g., Downing & Hamann (2013)) effects have been proposed for this common effect, though it should be noted that none of these explanations are specific to Mwera. Meanwhile, underlying /ND/ sequences are shortened, and the voiced consonant is deleted completely. Kenstowicz & Kisseberth (1977) and Noske, Schinkel and Smith (1982) point out that this effect can be captured simply via a counterfeeding order. If the deletion rule, turning ND sequences into N, applies before the voicing rule, then underlying /NT/ sequences will be ineligible for the deletion rule at the time that it applies.

**Language:** Nzɛbi

**ISO Code:** nzb

**Language Family:** Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Narrow Bantu, Northwest, B, Nzɛbi (B.52)

**Spoken in:** Gabon, Congo

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising,  $C \rightarrow D$ : Vowel Raising

**Domain:** Isomorphic

**Schematic:**  $a \rightarrow \varepsilon \rightarrow e \rightarrow i / \text{ɔ} \rightarrow o \rightarrow u$  (M)

**Environment(s):** /\_\_ {i}

**Source(s):** Guthrie (1968), Clements (1991), Flemming (1995) Kirchner (1996), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Kirchner (1996, p.344)*

$A \rightarrow B$ :  $a \rightarrow \varepsilon$   
/sal + i/  $\rightarrow$  [sɛli] ‘to work’  
 $B \rightarrow C$ :  $\varepsilon \rightarrow e, \text{ɔ} \rightarrow o$   
/βɛɛd + i/  $\rightarrow$  [βeedi] ‘to give’                      /tɔɔd + i/  $\rightarrow$  [toodi] ‘to arrive’  
 $C \rightarrow D$ :  $e \rightarrow i, o \rightarrow u$   
/bet + i/  $\rightarrow$  [biti] ‘to carry’                      /kolɔn + i/  $\rightarrow$  [kulini] ‘to go down’

**Description:** In Nzɛbi, several suffixes condition the raising of any vowel by one step. This creates a symmetrical system of back and front vowel raising, with an additional /a/  $\rightarrow$  [ɛ] step in the front portion of the chain. Kirchner (1996, p.344) points out that whilst the {-i} suffix in the above examples could suggest a phonological, harmonic motivation for the shift, the realization of the suffix is optional, and other, non-high suffixes also condition the raising. Despite this, Kirchner employs a phonological analysis of the effect, showing how conjunctions of PARSE constraints, working against a gradient RAISING constraint that assigns a violation for each step away from a low vowel a given output is, can model the shift. An alternative OT account, based on contrast, is postulated by Flemming (1995).

**Language:** Nzema

**ISO Code:** nzi

**Language Family:** Niger-Congo, Atlantic-Congo, Volta-Congo, Kwa, Nyo, Potou-Tano, Tano, Central, Bia, Southern

**Spoken in:** Ghana, Cote d'Ivoire

**Process(es) Involved:**  $A \rightarrow B$ : Voicing  
 $B \rightarrow C$ : Nasalization

**Domain:** Isomorphic

**Schematic:**  $nt \rightarrow nd \rightarrow nn$  (M)

**Environment(s):** / [+nas]\_\_

**Source(s):** Clopper (2001), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Clopper (2001, p.4)*

$A \rightarrow B$ :  $nt \rightarrow nd$   
/ɔ+n+tia/  $\rightarrow$  [ondia] 'he does not walk' (negative present)

$B \rightarrow C$ :  $nd \rightarrow nn$   
/ɔ+n+di/  $\rightarrow$  [onni] 'he does not eat' (negative present)

**Description:** In Nzema, the negative present and progressive tenses are marked by a process that Clopper explicitly defines as a chain shift (Clopper 2001, p.12). When the two-part prefix /ɔ+n/ is added to a stem beginning with a voiceless obstruent (the pattern is more widespread than Moreton's schematic suggests, with labial and velar plosives and coronal fricatives also undergoing the change), this consonant assimilates to the voiced character of the preceding nasal. The same prefix causes total nasal assimilation when the base form begins with a voiced consonant. It is suggested, though not made explicit, in Clopper's representations that this creates consonant sequences. Clopper uses local conjunction of IDENTIO[voice] and IDENTIO[nasal] to model this effect in Optimality Theory.

**Language:** Ojibwa

**ISO Code:** oji

**Language Family:** Algie, Algonquian, Ojibwa Potawatomi

**Spoken in:** Ontario, Canada

**Process(es) Involved:**  $A \rightarrow B$ : Place Assimilation  
 $B \rightarrow C$ : Unknown

**Domain:** N/A

**Schematic:**  $nk \rightarrow \eta \rightarrow \text{unknown}$  (M)

**Environment(s):** /\_\_k

**Source(s):** Kaye (1974), McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *example from McCarthy (1999, p.341)*

$A \rightarrow B$ :  $nk \rightarrow \eta$   
/takossin + k/  $\rightarrow$  [takoššinj] ‘(if) he arrives’

$B \rightarrow C$ :  $\eta \rightarrow \text{unknown}$   
No examples

**Description:** In Ojibwa, /n/ assimilates to the velar [ŋ] before /k/. This /k/ is subsequently deleted, both parts of which are represented by Moreton in the  $A \rightarrow B$  part of his schematic (as above). This is certainly an opaque generalization, as the reason for the velarization of what is known to be an underlyingly alveolar segment is not apparent from the surface representation. However, this is a *counterbleeding* interaction, as opposed to the kind of *counterfeeding* interactions present in the rest of the corpus. The  $B \rightarrow C$  part of the shift is motivated by Richness of the Base, the OT principle stating that “no constraints hold at the level of underlying forms” (Kager 1999, p.19). This presupposes a) that underlying /ŋ/ is possible in Ojibwa and that b) as the process above is the only way of getting surface [ŋ] in Ojibwa, these underlying /ŋ/s must map to something else. It is, however, impossible to say what this something is.

**Language:** Palauan

**ISO Code:** pau

**Language Family:** Austronesian, Malayo-Polynesian, Palauan

**Spoken in:** Palau

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Reduction (Shortening)  
 $B \rightarrow C$ : Vowel Reduction (Reduction to schwa/Deletion)

**Domain:** Isomorphic

**Schematic:**  $u: \rightarrow u \rightarrow \#/\emptyset$  (adapted slightly from M)

**Environment(s):** Unstressed syllables

**Source(s):** Flora (1974), Josephs (1975), Zuraw (2002), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Zuraw (2002, pp.2-4). NB: Underlying /uu/, in Zuraw's terms, is realized as [uw] on the surface*

$A \rightarrow B$ : /u:/  $\rightarrow$  [u]  
/búwʔə/  $\rightarrow$  [bu-él] 'betel nut'

$B \rightarrow C$ : /u/  $\rightarrow$  #, /u/  $\rightarrow$  [ə]  
/súbəð/  $\rightarrow$  [spəð-él] 'announcement'  
/kúk/  $\rightarrow$  [kək-úl] 'nail'

**Description:** Palauan has a complex system of vowel reduction which Zuraw explicitly refers to as a “synchronic chain shift” (2002, p. 1). Long vowels become short when they are no longer stressed (the examples above show that a stressed suffix can cause this effect), and short vowels in the same environment *either* apococate completely or are realized as schwa. It is the environment for the change (stressless syllables) and the  $u \rightarrow \emptyset$  mapping that explicitly mark out this effect as vowel reduction, particularly since Zuraw acknowledges that a previous survey of vowel reduction effects (Crosswhite, 1999) does not address complete vowel deletion.

**Language:** Pero

**ISO Code:** pip

**Language Family:** Afro-Asiatic, Chadic, West, A, A.2, Tangale, Tangale Proper

**Spoken in:** Various regions, Nigeria

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Lowering  
 $B \rightarrow C$ : Vowel Lowering

**Domain:** Isomorphic

**Schematic:**  $i \rightarrow e \rightarrow a$  (M)

**Environment(s):** / V(C)\_\_\_CC

**Source(s):** Frajzyngier (1989), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Frajzyngier (1989, p.41)*

$A \rightarrow B$ : /i/  $\rightarrow$  [e]  
/ní ɪll + kò/  $\rightarrow$  [niéllòyò] ‘I stood up’  
/cì mà + yì + n + nò/  $\rightarrow$  [cìmàyénnò] ‘if you (f) don’t make for me’

$B \rightarrow C$ : /e/  $\rightarrow$  [a]  
/àn + céngèw/  $\rightarrow$  [ánjáŋgèw] ‘one who is stubborn’  
/nì + céyy + kò/  $\rightarrow$  [nijáyórò] ‘I drunk all of it’

**Description:** In Frajzyngier’s analysis of Pero (1989), there are several lowering rules, most of which are one-step assimilation processes affecting the high vowels /i/ and /u/. However, there is one, highly circumscribed, circumstance in which both high and mid front vowels lower one step, in the pattern /i/  $\rightarrow$  [e]  $\rightarrow$  [a], a complete reversal of the standard raising chain shift pattern (*see e.g.*, Lena Spanish, Basaa, Nzébi). Frajzyngier himself specifies “that the rule applies only once, so there is no passage /i/  $\rightarrow$  [e]  $\rightarrow$  [a]” (p.40), a succinct summary of vowel-based shifts. Additionally, his somewhat unorthodox use of a multi-valent rule to model the process mirrors other rule-based analyses of chain shifts (for example Labov (1994) on diachronic shifts, see chapter 2 and chapter 5 for additional discussion).

**Language:** Pipil

**ISO Code:** ppl

**Language Family:** Uto-Aztecan, Southern Uto-Aztecan, Corachol-Aztecan, Core Nahua

**Spoken in:** Ocotepeque department, Dolores municipality, Honduras

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Devocalization

**Domain:** Adjacent

**Schematic:**  $VwV\{C/\#\} \rightarrow Vw\{C/\#\} \rightarrow Vh\{C/\#\}$  (M)

**Environment(s):** /\_\_ $\{C/\#\}$

**Source(s):** Campbell (1985), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Campbell (1985, pp.34-35)*

$A \rightarrow B$ :  $VwV\{C/\#\} \rightarrow Vw\{C/\#\}$   
/chikiwi-t/  $\rightarrow$  [nu-chikiw] ‘my basket’  
/kakawa-t/  $\rightarrow$  [nu-kakaw] ‘my cacao’

$B \rightarrow C$ :  $Vw\{C/\#\} \rightarrow Vh\{C/\#\}$   
/kuwa/ ‘to buy’  $\rightarrow$  [kuh-ki] ‘bought’  
/suwa/ ‘to spread out’  $\rightarrow$  [suh-ki] ‘spread’

**Description:** In the Nicaraguan language Pipil, and in its mother language Cuisnahut, there is a general process by which syllable-final /w/ loses its voicing and supralaryngeal features to become [h]. However, as Campbell notes “this rule does not apply to w in Cuisnahut which has come to be in the final position due to the application of the vowel loss rule in nouns” (1985, pp.34-35). These final /w/ segments are realized faithfully as [w]. Campbell suggests that the reason for this distribution lies in a counterfeeding order, wherein the devocalization rule changing /w/  $\rightarrow$  [h] occurs *before* the rule deleting final vowels, meaning that the context for the devocalization rule would no longer exist.

**Language:** Polish

**ISO Code:** pol

**Language Family:** Indo-European, Balto-Slavic, Slavic, West, Lechitic

**Spoken in:** Poland, Czech Republic, Germany

**Process(es) Involved:**  $A \rightarrow B$ : Palatalization  
 $B \rightarrow C$ : Palatalization

**Domain:** Isomorphic

**Schematic:**  $x \rightarrow \check{s} \rightarrow \acute{s}$  (adapted from M)

**Environment(s):** /\_\_i, e, j

**Source(s):** Rubach (1984), Łubowicz (2003b), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from (Łubowicz 2003b, p.315)*

$A \rightarrow B$ :  $x \rightarrow \check{s}$

nom. sg.	aug.	dimin.	gloss
gro[x]	gro[š]+ysk+o	gro[š]+ek	‘bean’
gma[x]	gma[š] + ysk + o	gma[š] + ek	‘building’
fartu[x]	fartu[š] + ysk + o	fartu[š] + ek	‘apron’

$B \rightarrow C$ :  $\check{s} \rightarrow \acute{s}$

gro[š]	gro[ś] + isk + o	gro[ś] + ik	‘a penny’
kapelu[š]	kapelu[ś] + ysk + o	kapelu[ś] + ik	‘hat’
arku[š]	arku[ś] + ysk + o	arku[ś] + ik	‘sheet’

**Description:** Polish has (at least) two distinct palatalization rules (following Rubach 1984). In the first, First Velar Palatalization, velar obstruents palatalize to their palatoalveolar counterparts before /i e j/. In the other, Nominal Strident Palatalization, underlying palatoalveolar /š/ is realized on the surface as prepalatal [ś]. Crucially, instances of ś that are derived from underlying /x/ do not undergo Nominal Strident Palatalization, as in the examples above. This creates, among many other forms, the near minimal pair *grošysko* ‘bean’ (aug.) vs. *grośisko* ‘a penny’ (aug.). Łubowicz calls this a chain shift and models it using Contrast Preservation constraints (2003b).

**Language:** Polish (Kashubian dialect)

**ISO Code:** csb

**Language Family:** Indo-European, Balto-Slavic, Slavic, West, Lechitic

**Spoken in:** Pomeranian province, Poland

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** N/A

**Schematic:**  $a \rightarrow e \rightarrow \text{ɪ}$

**Environment(s):** None given by Łubowicz

**Source(s):** Łubowicz (2003a), Sanders (2003)

**Examples:** *All examples are from Łubowicz (2003a, pp.48-49) and show differences between standard Polish and Kashubian Polish*

$A \rightarrow B$ :  $a \rightarrow e$   
[trava] vs. [treva] ‘grass’  
[gada] vs. [gede] ‘he talks’

$B \rightarrow C$ :  $e \rightarrow \text{ɪ}$   
[ser] vs. [sɪr] ‘cheese’  
[mleko] vs. [mlɪko] ‘milk’

**Description:** This shift is discussed by Łubowicz (2003a) as an example of a two-step raising shift. The shift appears to be fairly general and non-assimilatory in nature, however there are complexities, for example the difference between the raising of both vowels in [gada]  $\rightarrow$  [gede] as opposed to [trava]  $\rightarrow$  [treva]. Łubowicz does not address this particular concern, but she does discuss the “incomplete” (p.52) nature of the shift, in that the raising of /e/ frequently only reaches [ɪ] as opposed to the perhaps more expected [i] (*see e.g.*, Lena Spanish, Basaa etc). Indeed, every part of the chain is subject to this incompleteness, as she notes that “In some regions of Kashubian...the low vowel raises but it does not make it all the way to [e]” (p.48, note 16).

**Language:** Sanskrit

**ISO Code:** san

**Language Family:** Indo-European, Indo-Iranian, Indo-Aryan

**Spoken in:** Various regions, India

**Process(es) Involved:**  $A \rightarrow B$ : Coalescence  
 $B \rightarrow C$ : Unknown

**Domain:** N/A

**Schematic:** ai/au  $\rightarrow$  e:/o:  $\rightarrow$  unknown (M)

**Environment(s):** /C\_\_C

**Source(s):** Gonda (1966), Gnanadesikan (1997), Moreton (2004a)

**Examples:** *all examples from Gnanadesikan (1997, p.140)*

$A \rightarrow B$ : ai  $\rightarrow$  e:

ca + ihi  $\rightarrow$  ceeha ‘and here’

tvaa + iiś vara  $\rightarrow$  tveeś vara ‘you, O Lord’

$B \rightarrow C$ : au  $\rightarrow$  o:

ca + uktam  $\rightarrow$  cooktam ‘and said’

saa + uvacaa  $\rightarrow$  soovaaca ‘she said’

**Description:** In Sanskrit, mid vowels are only present as a result of a process of coalescence; there are no underlying mid-vowel. This means that labelling the coalescence process a chain shift is a purely theory-internal assumption, only possible under strict interpretations of Optimality Theory. As Gnanadesikan notes, “[w]ithin Optimality Theory, the constraints are all on the output, or on the relationship between the input and the output. It is therefore impossible to rule out mid vowels in the input” (1997, p.141). Moreton’s logic thus becomes clear: underlying diphthongs become long mid vowels. Because of Richness of the Base, one has to assume that there may be underlying mid vowels in Sanskrit. This is why the third stage of the putative shift is unknown (see 5.2.3.1 for more discussion).

**Language:** Sea Dayak

**ISO Code:** iba

**Language Family:** Austronesian, Malayo-Polynesian, Malayo-Chamic, Malayic, Ibanic

**Spoken in:** Various regions, Malaysia

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Nasal Harmony

**Domain:** Overlapping

**Schematic:**  $\eta ga \rightarrow \eta a \rightarrow \eta \tilde{a}$  (adapted slightly M)

**Environment(s):**  $A \rightarrow B$ :  $/[+nas, +cons]_\_$ ,  $B \rightarrow C$   $/[+nas]_\_$

**Source(s):** Scott (1957), Kenstowicz & Kisseberth (1979), Moreton & Smolensky (2002), Moreton (2004a), Łubowicz (2011)

**Examples:** *all examples from Łubowicz (2011, section 3.1)*

$A \rightarrow B$ :  $\eta ga \rightarrow \eta a$   
 $/na\eta ga/ \rightarrow [n\tilde{a}\eta a?]$  ‘set up a ladder’

$B \rightarrow C$ :  $\eta a \rightarrow \eta \tilde{a}$   
 $/na\eta a/ \rightarrow [n\tilde{a}\eta \tilde{a}?]$  ‘straighten’

**Description:** Sea Dayak has a productive process of nasal harmony, discussed in Kenstowicz & Kisseberth (1979). Vowels following nasal consonants are nasalized. Whilst this harmony can spread to multiple vowels, and certain consonants, such as glides, are transparent to it, the process is blocked by obstruents. Sea Dayak also contains a cluster simplification process (called ‘Nasal Cluster Simplification’ by Kenstowicz & Kisseberth) whereby voiced obstruents are deleted when they directly follow a nasal. The interaction of these two processes explains the seeming underapplication of nasal harmony in the form  $[n\tilde{a}\eta a?]$ , ‘set up a ladder’. As the final  $[a]$  directly follows a nasal, the expectation would be that nasal harmony should occur. Indeed, it is only this lack of nasality that marks this form as distinct from  $[n\tilde{a}\eta \tilde{a}?]$ , ‘straighten’. See 5.3.2 for much more discussion.

**Language:** Serbo-Croatian 1

**ISO Code:** hbs

**Language Family:** Indo-European, Balto-Slavic, Slavic, South, Western

**Spoken in:** Bosnia-Herzegovina, Croatia, Serbia

**Process(es) Involved:**  $A \rightarrow B$ : Vocalization  
 $B \rightarrow C$ : Stress Assignment

**Domain:** Non-overlapping

**Schematic:**  $V(C)l\# \rightarrow 'V(C)o\# \rightarrow V(C)'o\#$  (M)

**Environment(s):**  $A \rightarrow B$ : Final syllable  $B \rightarrow C$ :  $\_\_C_{(0)}\#$

**Source(s):** Kenstowicz (1994), Moreton (2004a)

**Examples:** *all examples from Kenstowicz (1994, pp.90-91)*

$A \rightarrow B$ : $V(C)l\# \rightarrow V'(C)o\#$		
Masc	Fem	Gloss
debéo	debelá	'fat'
béo	belá	'white'

$B \rightarrow C$ : $'V(C)o\# \rightarrow V(C)'o\#$	
/dobro/ $\rightarrow$ [dobró]	'good' (neut)
/jasno/ $\rightarrow$ [jasnó]	'clear'

**Description:** Kenstowicz (1994, p. 90) postulates a general rule that in monosyllabic or disyllabic Serbo-Croatian adjectives, the suffix that is added to specify gender or number takes stress. This rule takes the form  $V \rightarrow [+stress] / \_\_C_{(0)}\#$ , meaning that final vowels take stress in these forms. This is the process illustrated in the  $B \rightarrow C$  examples above. However, there are certain exceptions to this rule, as shown in the  $A \rightarrow B$  examples [debéo] and [béo], where stress is penultimate (see the other Serbo-Croatian entry in this database for another class of exceptions). Kenstowicz suggests that these forms have an underlying final /l/ which is vocalized to [o] word-finally, but that this vocalization only happens *after* the stress rule has applied. This counterfeeding order means that when the stress rule applies to the underlying form /bel/, the final vowel is /e/. By the time /bel/ becomes [béo], stress has already been applied and cannot shift to the final vowel.

**Language:** Serbo-Croatian 2

**ISO Code:** hbs

**Spoken in:** Bosnia-Herzegovina, Croatia, Serbia

**Language Family:** Indo-European, Balto-Slavic, Slavic, South, Western

**Process(es) Involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Stress Assignment

**Domain:** Non-overlapping

**Schematic:**  $VCC\# \rightarrow 'VCaC\# \rightarrow VC'aC\#$  (2<sup>nd</sup> C [+son]) (M)

**Environment(s):**  $A \rightarrow B$ : /C\_C,  $B \rightarrow C$ : Final syllables

**Source(s):** Kenstowicz (1994), Moreton (2004a)

**Examples:** *all examples from Kenstowicz (1994)*

$A \rightarrow B$ :  $VCC\# \rightarrow 'VCaC\#$   
/dobr/  $\rightarrow$  [dóbar] 'good' (masc)  
/jasn/  $\rightarrow$  [jásan] 'clear' (masc)

$B \rightarrow C$ :  $'VCaC\# \rightarrow VC'aC\#$   
n.a

**Description:** Kenstowicz discusses a rule of epenthesis in Serbo-Croatian, whereby word-final consonant clusters in masculine adjectival forms are broken up by the introduction of an epenthetic [a]. As is fairly common, epenthetic vowels cannot take stress (see the entries for Bduul Arabic and Mohawk in this chapter, Broselow (2008) or Gouskova & Hall (2009) for other examples). This creates a class of exceptions to the general rule of Serbo-Croatian adjectives that places stress word-finally (see the other putative Serbo-Croatian shift on the previous page for another class of exceptions). However, in this case Moreton's schematic is problematic. This is because, in the text that he cites (Kenstowicz, 1994, pp. 90-94), there are no examples of the C part of the shift. That is to say, there are no examples of a word which takes the form  $VC'aC\#$ .

**Language:** Shoshone (Tümpisa Panamint dialect)

**ISO Code:** par

**Language Family:** Uto-Aztecan, Northern Uto-Aztecan, Numic, Central

**Spoken in:** California, USA

**Process(es) Involved:**  $A \rightarrow B$ : Epenthesis  
 $B \rightarrow C$ : Consonant Deletion

**Domain:** Overlapping

**Schematic:**  $V \rightarrow V?V \rightarrow VV$  (M. Very doubtful, *Moreton's parentheses*)

**Environment(s):**  $A \rightarrow B$ : /\_\_#,  $B \rightarrow C$ : V\_\_V

**Source(s):** Dayley (1989), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Dayley (1989, pp.417-418)*

$A \rightarrow B$ :  $V \rightarrow V?V$

/tua'ʔ/  $\rightarrow$  [tua'a]  $\rightarrow$  [túaʔA] 'son'

/poto'o/  $\rightarrow$  [poto'o]  $\rightarrow$  [póroʔO] 'staff, cane'

$B \rightarrow C$ :  $V?V \rightarrow VV$

/mo'o/  $\rightarrow$  [mǽʔo] ~ [mǽo] 'hand'

/tühüya/  $\rightarrow$  [tìhiyà] ~ [tìyà] 'deer'

**Description:** Tümpisa Shoshone, a language with less than 10 speakers when recorded by Dayley in 1989, contains a process where “[a] glottal stop and voiceless echo vowel are often inserted at the end of words in phrase final position” (Dayley, 1989, p. 418). This represents the  $A \rightarrow B$  step in the putative chain shift. However, this is quite an unusual step as it adds two full segments to the underlying form. In a non-linear analysis, this could be analysed as a partially reduplicative template, as the echo vowel always has the character of the full vowel that precedes it. The second part of the putative shift is characterized by the loss of a glottal stop from between two vowels (though as the second example shows, any segment without supralaryngeal features may undergo this process).

**Language:** Southeastern Pomo

**ISO Code:** pom

**Language Family:** Pomoan

**Spoken in:** California, USA (only 1 native speaker as of 2014)

**Process(es) Involved:**  $A \rightarrow B$ : Metathesis  
 $B \rightarrow C$ : Vowel Lowering

**Domain:** Adjacent

**Schematic:**  $Ci \rightarrow iC \rightarrow eC$  (doubtful) (M. Moreton's parentheses)

**Environment(s):**  $A \rightarrow B$ :  $\_\_C$ ,  $B \rightarrow C$ :  $\_\_C$

**Source(s):** Moshinsky (1974), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Moshinsky (1974, pp.29-31)*

$A \rightarrow B$ :  $Ci \rightarrow iC$

/ʔson+k/  $\rightarrow$  ʔesón+ki  $\rightarrow$  [ʔesónik] 'to guess'

/ʔteé + mku/  $\rightarrow$  [ʔetéémuk] 'twice'

/t'o+m+c/  $\rightarrow$  t'ó+m+ci  $\rightarrow$  [t'ómić] 'razor'

$B \rightarrow C$ :  $iC \rightarrow eC$

/ʔwal + é + l + t/  $\rightarrow$  ʔuwal + ci + li + t  $\rightarrow$  [ʔuwálcelit] 'he ducked repeatedly'

/ca + mlu + l + t/  $\rightarrow$  [cámlolit] 'he ran around'

/da.fʔey + é + t/  $\rightarrow$  da fʔéy + ci + t  $\rightarrow$  [da fʔéyćes] 'he didn't bother you'

/ci + mku + é + t/  $\rightarrow$  [címkoćit] 'those three are fighting each other'

**Description:** Moshinsky (1974) postulates a rule of word-final vowel metathesis, which causes word-final vowels in verbs or deverbal nouns to surface prior to the final consonant of the word. Products of this metathesis rule fail to undergo an otherwise productive and obligatory process of vowel lowering. Both *i* and *u* lower before *l*, with *i* additionally lowering before *s* and *u* lowering before *ć* (p.28-29). All of the examples of *i* illustrated in the above examples are epenthetic. This epenthesis operates either morpheme-finally or in three-consonant clusters, in which case the epenthetic vowel (always *i* except before {m, q}, where *a* surfaces) is inserted between C2 and C3 (1974, p.28). Moshinsky gives an explicit order for all of the rules of the language (p.18), in which epenthesis is ordered before vowel lowering and metathesis.

**Language:** Spanish (Gran Canaria dialect, also known as North Corsican)

**ISO Code:** spa

**Language Family:** Indo-European, Italic, Romance, Italo-Western, Western, Gallo-Iberian, Ibero-Romance, West Iberian, Castilian

**Spoken in:** Gran Canaria

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Lenition (Voicing)  
 $B \rightarrow C$ : Consonant Lenition (Spirantization)

**Domain:** Isomorphic

**Schematic:** p t k  $\rightarrow$  b d g  $\rightarrow$   $\beta$   $\delta$   $\gamma$

**Environment(s):** /V\_\_V

**Source(s):** Oftedal (1985) Gussenhoven & Jacobs (2011), Broś (2015)

**Examples:** *all examples from Gussenhoven & Jacobs (2011, p.118)*

$A \rightarrow B$ : p t k  $\rightarrow$  b d g  
/tipiko/  $\rightarrow$  [tibigo] ‘typical’  
/la kama/  $\rightarrow$  [la gama] ‘the bed’  
/una tjenda/  $\rightarrow$  [una djenda] ‘a shop’

$B \rightarrow C$ : b d g  $\rightarrow$   $\beta$   $\delta$   $\gamma$   
/roba/  $\rightarrow$  [ro $\beta$ a] ‘he/she steals’  
/nada/  $\rightarrow$  [na $\delta$ a] ‘nothing’  
/la gana/  $\rightarrow$  [la  $\gamma$ ana] ‘the appetite’

**Description:** In the Gran Canaria dialect of Spanish, there is a consistent process of intervocalic lenition. The exact kind of lenition depends on the nature of the segment that is undergoing the process. Underlyingly voiceless plosives become voiced, as in the  $A \rightarrow B$  examples above, and underlyingly voiced plosives spirantize, as shown in the  $B \rightarrow C$  examples. Crucially, underlying voiceless plosives do not spirantize. Gussenhoven & Jacobs (2011) label the process a counterfeeding effect, though Broś (2015) is explicit that she does consider the process to be a chain shift. Gussenhoven & Jacobs analyse the effect using four different versions of Optimality Theory, concluding that Optimality Theory with Candidate Chains (see McCarthy(2007)) is the best method for dealing with the effect.

**Language:** Spanish (Lena dialect)

**ISO Code:** spa

**Spoken in:** Lena municipality, Asturias region, Spain

**Language Family:** Indo-European, Italic, Romance, Italo-Western, Western, Gallo-Iberian, Ibero-Romance, West Iberian, Castilian

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Isomorphic

**Schematic:**  $a \rightarrow e \rightarrow i$  (NB: In Moreton's corpus, the shift is incorrectly schematized as  $i \rightarrow e \rightarrow a$ )

**Environment(s):** /\_\_{-u}

**Source(s):** Hualde (1989), Parkinson (1996), Gnanadesikan (1997), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Parkinson (1996, p.19)*

$A \rightarrow B$  :  $a \rightarrow e$

[gáta] 'cat' (fem sg.) vs. [gétu] 'cat' (masc sg.)

[sánta] 'saint' (fem sg.) vs. [séntu] 'saint' (masc sg.)

$B \rightarrow C$ :  $e \rightarrow i$

[néna] 'child' (fem sg.) vs. [nínu] 'child' (masc sg.)

[bwéna] 'good' (fem sg.) vs. [bwínu] 'good' (masc sg.)

**Description:** A regular vowel raising effect is observed in the Lena dialect of Spanish. This is fairly clearly a process of assimilation, as the raising is conditioned by the presence of the high back vowel /u/ in the masculine singular form. This effect is part of the larger set of metaphony processes common in Romance languages (for discussion see Calabrese (2011)), but whilst these effects usually show only a raising from mid to high vowels, this shift has three parts, from low to mid to high. The effect is partially symmetrical, with back mid vowels rising to high vowels in the same situation, for example [bóna] 'good' (fem sg.)  $\rightarrow$  [búnu] 'good' (masc sg.) (Parkinson (1996, p.19)), but the low vowel /a/ only raises to [e] and never [o].

**Language:** Tarascan

**ISO Code:** tsz

**Language Family:** Tarascan

**Spoken in:** Various regions, Mexico

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Vowel Deletion

**Domain:** Adjacent

**Schematic:**  $ae \rightarrow ee \rightarrow e$  (M)

**Environment(s):** /\_\_e

**Source(s):** Foster (1968), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** (Note: I have edited Foster's original representations for clarity. There should not be any difference in meaning.)

$A \rightarrow B$ :  $ae \rightarrow ee$

/e-re-ta-eča/  $\rightarrow$  [eréteeča] 'towns' (Foster, 1968, p. 36)

/i-má-eri/  $\rightarrow$  [iméeri] 'his/her/its of that one' (Foster, 1968, p. 73)

$B \rightarrow C$ :  $ee \rightarrow e$

/a - mpe - e - ka - ni/  $\rightarrow$  /ampékani/ 'to be what?' (Foster, 1968, p. 36)

/i-nte-eri/  $\rightarrow$  [intéri] 'his/her/its of that one' (Foster, 1968, p. 73)

**Description:** These two rules are introduced without any discussion in Foster's grammar of Tarascan (1968). The first rule appears to be a straightforward assimilation process, of which Tarascan has many (see Foster 1968, p.35). It appears that the mid vowel is rather unstable, as it is the target for a wide range of phonological processes, including that captured by the second part of the putative shift. In this process, when a morpheme ending with /e/ is concatenated with another morpheme beginning with /e/, the resulting /e-e/ sequence is shortened to surface [e]. This does not occur in sequences derived by the first rule. A problem for a chain shift analysis is that sometimes complete neutralisation, or /A/  $\rightarrow$  /C/ patterning, occurs, where an /ae/ sequence becomes [e], as opposed to the expected [ee] (see Foster, p.50): still more problematically, in one case /ae/  $\rightarrow$   $\emptyset$  (see Foster, p.59).

**Language:** Thok Reel

**ISO Code:** atu

**Language Family:** Nilo-Saharan, Eastern Sudanic, Nilotic, Western, Dinka-Nuer, Nuer

**Spoken in:** Various regions, South Sudan

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Lowering  
 $B \rightarrow C$ : Vowel Lowering  
 $C \rightarrow D$ : Vowel Lowering

**Domain:** Isomorphic

**Schematic:**  $\underline{e} \rightarrow e \rightarrow \varepsilon \rightarrow a$  (*note*: underlining denotes breathiness)

**Environment(s):** All present tense conjugations *except* 2<sup>nd</sup>/3<sup>rd</sup> person singular

**Source(s):** Reid (2010), Trommer (2011)

**Examples:** *all examples from Reid (2010, p.78)*

$A \rightarrow B$ :  $\underline{e} \rightarrow e$

[k $\underline{e}$ eer] ‘awaken someone’ (3 sg.) vs. [k $\underline{e}$ eer-kɔ(n)] ‘awaken someone’ (1 sg.)  
[p $\underline{e}$ een] ‘prevent’ (3 sg.) vs. [p $\underline{e}$ een-kɔ(n)] ‘prevent’ (1 sg.)

$B \rightarrow C$ :  $e \rightarrow \varepsilon$

[teT] ‘dig’ (3 sg.) vs. [t $\varepsilon$ t-kɔ(n)] ‘dig’ (1 sg.)  
[kw $\underline{e}$ en] ‘count’ (3 sg.) vs. [kw $\varepsilon$ ’en-kɔ(n)] ‘count’ (1 sg.)

$C \rightarrow D$ :  $\varepsilon \rightarrow a$

[ $\eta$  $\varepsilon$ C] ‘know’ (3 sg.) vs. [ $\eta$ aC-kɔ(n)] ‘know’ (1 sg.)

**Description:** There is a pervasive process of vowel lowering in Thok Reel, which occurs in every conjugation of the present tense except the 2<sup>nd</sup> and 3<sup>rd</sup> person singular forms. Trommer discusses the effect above as a chain shift, given that breathy / $\underline{e}$ / surfaces as modal [e], whilst underlying modal /e/ surfaces as [ε] and underlying / $\varepsilon$ / surfaces as [a]. It should, however, be pointed out, that there are other mappings that form part of the Thok Reel lowering effect that do not form a chain shift. High vowels, for example, lower and diphthongize, for example /i/  $\rightarrow$  [jε]. Thok Reel does not have phonemic diphthongs, so there is no further movement, and thus no chain shift (see 5.4.2.2 for more discussion).

**Language:** Toba Batak

**ISO Code:** bbc

**Language Family:** Austronesian, Malayo-Polynesian, Northwest Sumatra-Barrier Islands, Batak, Southern

**Spoken in:** Indonesia, Sumatra

**Process(es) Involved:**  $A \rightarrow B$ : Assimilation (Denasalization)  
 $B \rightarrow C$ : Glottalization

**Domain:** Isomorphic

**Schematic:**  $NT \rightarrow TT \rightarrow ?T$  (M)

**Environment(s):** / \_\_[-son, -cont, -voi]

**Source(s):** Hayes (1986), Pater (1999), Moreton and Smolensky (2002), Moreton (2004a)

**Examples:** *all examples from Hayes (1986, pp.480-481)*

$A \rightarrow B$ :  $NT \rightarrow TT$

/baoa an pɛddek/ → [baoa ap pɛddek] 'that man is short'

$B \rightarrow C$ :  $TT \rightarrow ?T$

/lap piŋgɔl/ → [laʔ piŋgɔl] 'wipe off the ear'

/halak batak/ → [halaʔ batak] 'Batak person'

**Description:** The putative shift in Toba Batak rests on the interaction of a process of denasalization of nasal stops before voiceless consonants ( $NT \rightarrow TT$ ) and glottalization, in which the first member of a homorganic TT cluster becomes a glottal stop. Hayes (1986) notes that the TT clusters created by denasalization share a place and manner of articulation, but that glottalization does not occur in these clusters. Hayes, posits that the result of denasalization is that the new homorganic TT clusters share a doubly linked [-voice] feature, making them 'true' geminates, thus inalterable. Under this kind of analysis, any process that acts on a geminate must act on the entire segment, as opposed to the first or second 'half' alone. Therefore, glottalization cannot apply. In the  $TT \rightarrow ?T$  examples, however, no features are genuinely shared, making a dissimilation process such as glottalization a likely outcome.

**Language:** Tzutujil (Santiago dialect)

**ISO Code:** tzj

**Language Family:** Mayan, Yucatecan-Core Mayan, K'ichean-Mamean, K'ichean, Poqom-K'ichean, Core K'ichean, Kaqchikel-Tz'utujil

**Spoken in:** Various regions, Guatemala

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Vowel Shortening

**Domain:** Adjacent

**Schematic:**  $V?C^2 \rightarrow V:C^2 \rightarrow VC^2$  (M)

**Environment(s):**  $A \rightarrow B$ :  $/V\_C$ ,  $B \rightarrow C$ :  $/\_C$

**Source(s):** Dayley (1985), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:**

$A \rightarrow B$ :  $V?C^2 \rightarrow V:C^2$   
 $/xch'o?b'a/ \rightarrow [xch'oob'a]$  'it was thought'  
 $/xyit?tz'a/ \rightarrow [xyiit'za]$  'it was squeezed' (Dayley, 1985, p. 50)

$B \rightarrow C$ :  $V:C^2 \rightarrow VC^2$   
[b'iix] 'song' vs. [b'ixaniem] 'to sing'  
[muuj] 'shadow, shade' vs. [nmujaal] 'my shadow' (Dayley, 1985, p. 45)

**Description:** The basic pattern in the  $A \rightarrow B$  part of the shift is that glottal stops are deleted before a consonant, (presumably) causing compensatory lengthening in the preceding vowel. However, there is also a more general conflicting rule in which long vowels are shortened before a consonant, which does not apply to forms derived by the previous rule. The schematic, reproduced from Moreton's online corpus, may be somewhat misleading as it suggests that the environment for both parts of the shift is before a glottalised consonant. Whilst Dayley specifies that this is a necessary environment for the first rule in the example set (corresponding to the  $/A/ \rightarrow [B]$  part of the putative shift), the more general rule does not have such a restriction on it.

**Language:** Wintu 1

**ISO Code:** wnw

**Language Family:** Wintuan

**Spoken in:** Northern California, USA (no known living speakers)

**Process(es) Involved:**  $A \rightarrow B$ :  $r \rightarrow \theta$  (Non-leniting Spirantization)  
 $B \rightarrow C$ : Unknown

**Domain:** N/A

**Schematic:**  $rh \rightarrow \theta \rightarrow$  unknown (M)

**Environment(s):**  $A \rightarrow B$ :  $\_\_\#$ ,  $B \rightarrow C$ : unknown

**Source(s):** Pitkin (1984), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *from Pitkin (1984, p.47)*

$A \rightarrow B$ :  $rh \rightarrow \theta$   
/nurh/  $\rightarrow$  [nu $\theta$ ]  
BUT  
/nurhdoli/  $\rightarrow$  [nuhdoli]

$B \rightarrow C$ : unknown

**Description:** In Wintu, there is productive spirantisation and reduction of word-final sonorant-h clusters, as in the first example above. /h/ is deleted and /r/ spirantizes to [θ]. However, /rh/ clusters word-medially do not behave the same way as they do word-finally, as is shown in the second example. The picture is fairly complex. First, Pitkin states that this effect is “a unique simplification of underlying /rhd/” (1984, p. 47), which suggests that, if followed by a vowel, /rh/ clusters will also act as they do word-finally. It may be that there is a constraint against fricative-stop sequences that forces a reduction to /h/ and blocks the spirantisation. If this is the case, then there is not a chain shift effect, simply word-position effects. Before a vowel or the end of a word, /rh/ clusters become [θ], and before a consonant they become /h/. This is a counterbleeding order, as the environment for spirantisation is deleted after it has applied.

**Language:** Wintu 2

**ISO Code:** wnw

**Language Family:** Wintuan

**Spoken in:** Northern California, USA (no known living speakers)

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Vowel Raising

**Domain:** Overlapping

**Schematic:**  $\epsilon C C a \rightarrow \epsilon C a \rightarrow i C a$  (M)

**Environment(s):**  $A \rightarrow B$ : /C \_\_,  $B \rightarrow C$ : /\_\_ Ca

**Source(s):** Pitkin (1984), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:**

$A \rightarrow B$ :  $\epsilon C C a \rightarrow \epsilon C a$  (adapted from Pitkin (1984, p. 46))  
/ʔelewwar/  $\rightarrow$  [ʔelewar] ‘not’

$B \rightarrow C$ :  $\epsilon C a \rightarrow i C a$  (adapted from Pitkin (1984, p. 44))  
/lEla/  $\rightarrow$  [lila] ‘to transform’  
/dEka/  $\rightarrow$  [dika] ‘to climb’

**Description:** A regular process in Wintu is the reduction of homogenous consonant clusters, so that they are realized as singletons on the surface. In Moreton’s corpus, it is implied that this interacts with a separate vowel ablaut process in the language, in which mid vowels raise preceding a /Ca/ sequence (Pitkin (1984) proposes an explanation of dissimilation for this effect). Pitkin suggests that there is an ordering relation between the two processes; “[i]n particular, the root-deriving transitive suffix {c} induces clusters which, if simplified before application of the rules of vowel dissimilation, would often...lead to raising, where in fact such raising does not occur, since it is blocked by the consonant cluster” (p.46).

**Language:** Xiamen

**ISO Code:** nan

**Language Family:** Sino-Tibetan, Chinese

**Spoken in:** Various regions, Southern China, Taiwan

**Process(es) Involved:** *A* → *B*: Tone Sandhi  
*B* → *C*: Tone Sandhi  
*C* → *D*: Tone Sandhi  
*D* → *E*: Tone Sandhi  
*E* → *B*: Tone Sandhi

**Domain:** Isomorphic

**Schematic:** 24 → 22 → 21 → 53 → 44 → 22 (at which point everything bar italic step starts again)

**Environment(s):** Non-final

**Source(s):** W. Wang (1967), H.Hsieh (1976), S. Anderson (1978), Chen (1987), Moreton (2004b), F.Hsieh (2005) Barrie (2006), G. Thomas (2008)

**Examples:** *examples from Chen (1987), reproduced in Thomas (2008, p. 422)*

Citation Form:	Translation:	Sandhi Form:	Translation:
we-24	<i>shoe</i>	we-22 tua-21	<i>shoe laces</i>
wi-22	<i>stomach</i>	wi-21 pih-22	<i>stomach ache</i>
tsu-21	<i>house</i>	tsu-53 ting-53	<i>roof top</i>
hai-53	<i>ocean</i>	hai-44 kih-24	<i>ocean-front</i>
pang-44	<i>fragrant</i>	pang-22 tsui-53	<i>fragrant water</i>

**Description:** The much-discussed Xiamen tone circle takes the form of a sandhi effect. When pronounced in prominent positions (roughly speaking at the right edge of a phrase) or in isolation, words in Xiamen have a citation form which takes a particular tone. In other contexts, these words take separate sandhi tones. The shift from citation to sandhi tones takes the form of a circle, with the exception of the /24/ → [22] part of the shift, which creates a neutralization effect (as citation forms with both /24/ and /44/ take [22] as their sandhi tone).

**Language:** Yagua

**ISO Code:** yad

**Language Family:** Yaguan

**Spoken in:** Loreto region, Peru

**Process(es) Involved:**  $A \rightarrow B$ : Consonant Deletion  
 $B \rightarrow C$ : Vowel Assimilation (Height)

**Domain:** Overlapping

**Schematic:**  $V_1hV_2 \rightarrow V_1V_2 \rightarrow V_2V_2$  (M)

**Environment(s):**  $A \rightarrow B$ : /V\_\_V#,  $B \rightarrow C$ : /\_\_V

**Source(s):** Payne & Payne (1990), Moreton (2004a)

**Examples:**

$A \rightarrow B$ :  $V_1hV_2 \rightarrow V_1V_2$  (examples adapted from Payne & Payne 1990, p.439, 440)

/sa-suuta-janũ/  $\rightarrow$  [sasuutããñũ] ‘He washed long ago’

/riy-pichu-janu/  $\rightarrow$  [ripichõõñũ] ‘They tied (it) long ago’

$B \rightarrow C$ :  $V_1V_2 \rightarrow V_2V_2$  (examples adapted from Payne & Payne 1990, p.442)

/taryũũy-ta-i/  $\rightarrow$  [taryũũyĩĩ] ‘seller’

/sa-jũnããy/  $\rightarrow$  [sũnããy] ‘he cries’

**Description:** In the Peruvian language Yagua, underlying /h/ is productively deleted between two vowels at morpheme boundaries, creating two-vowel sequences.

Another general rule in Yagua is progressive vowel assimilation. This also occurs at morpheme boundaries. The final vowel of the first morpheme takes on the character of the second vowel. According to Payne & Payne “[t]he vowel assimilation rule is ordered before *j*-deletion, preventing its application to forms which had *j*’s [*j* is orthographic /h/] in their underlying form” (p.443). This is presumably the justification for Moreton’s schematic. Note, however, that forms with underlying /h/ undergo a separate process of coalescence if the vowels do not match in height. This also suggests an alternative hypothesis that is nothing to do with underlying /h/. In all of Payne & Payne’s examples, full assimilation occurs when the sequence is a low vowel followed by a high vowel, but coalescence occurs when the sequence is a high vowel followed by a low vowel. This would preclude a chain shift analysis, as the deletion of /h/ would cease to be relevant.

**Language:** Yokuts (Wikchamni dialect) 1

**ISO Code:** yok

**Language Family:** Yokutsan

**Spoken in:** San Joaquin Valley, California, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Lowering  
 $B \rightarrow C$ : Unknown

**Domain:** N/A

**Schematic:**  $i: \rightarrow e: \rightarrow$  Unknown (M)

**Environment(s):** Unconditioned

**Source(s):** Gamble (1978), McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:**

$A \rightarrow B$ :  $i: \rightarrow e$

/piyi + i:na/  $\rightarrow$  [piye:na] ‘will finish’ (Archangeli & Suzuki, 1997)

/k<sup>h</sup>ama + i:na/  $\rightarrow$  [k<sup>h</sup>ame:na] ‘will dance’ (Gamble 1978, p.17)

$B \rightarrow C$ :  $e: \rightarrow$  Unknown

No examples given

**Description:** In the Wikchamni dialect of Yokuts, there is a regular rule of lowering that states that long vowels that are underlyingly [+high] must surface as [-high].

This is usually claimed to be the only source of surface mid vowels in the language (the arguments for this are similar to those used to describe Yowlumne, in for example Kenstowicz & Kisseberth (1979)). If there is no underlying /e/ in the language then a chain shift analysis is not possible. A chain of  $A \rightarrow B \rightarrow C$  relies on the notion of underlying A mapping to surface B *whilst* underlying B maps to surface C. Therefore, all parts of the shift must be present underlyingly.

**Language:** Yokuts (Wikchamni dialect) 2

**ISO Code:** yok

**Language Family:** Yokutsan

**Spoken in:** San Joaquin Valley, California, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Raising  
 $B \rightarrow C$ : Rounding Harmony

**Domain:** Overlapping

**Schematic:**  $o \dots i \rightarrow u \dots i \rightarrow u \dots u$  (M)

**Environment(s):**  $A \rightarrow B$ : /\_\_C<sub>(0)</sub>i,  $B \rightarrow C$ : /\_\_C<sub>(0)</sub>u

**Source(s):** Archangeli & Suzuki (1997), McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:**

$A \rightarrow B$ :  $o \dots i \rightarrow u \dots i$  (from Archangeli & Suzuki 1997, p.218)

/t'oyx-in/  $\rightarrow$  [t'uyxin] 'will doctor'

/pot-k'in/  $\rightarrow$  [putk'in] 'will sour'

$B \rightarrow C$ :  $u \dots i \rightarrow u \dots u$  (from McCarthy 1999, p.41)

/luk'l-hin/  $\rightarrow$  [luk'ulhun] 'buries/might bury'

**Description:** In the Wikchamni dialect of Yokuts, there is an interaction of two processes that share an assimilatory nature but operate on different domains. The  $A \rightarrow B$  part of the putative shift is formed by a process of raising, whereby underlyingly mid back vowels assimilate in height to a following suffixal *i*. This interacts with a general process of rounding harmony, whereby suffixal *i* becomes surface [u] when preceded directly by *u* (ignoring intervening consonants). When *u* in the stem is derived by the height assimilation, *i* in the suffix fails to undergo rounding harmony. This could certainly be modeled by a counterfeeding order, though McCarthy (1999) uses sympathy theory in OT to model the putative shift.

**Language:** Yokuts (Yowlumne dialect) 1

**ISO Code:** yok

**Language Family:** Yokutsan

**Spoken in:** San Joaquin Valley, California, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Lowering  
 $B \rightarrow C$ : Vowel Lowering

**Domain:** Isomorphic

**Schematic:**  $i: \rightarrow e:/e \rightarrow a:/a$  (M)

**Environment(s):** Unconditioned

**Source(s):** Kenstowicz & Kisseberth (1979), Archangeli (1991), McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a), D'Arcy (2003)

**Examples:** *examples from D'Arcy (2003, p.29)*

$A \rightarrow B$ :  $i: \rightarrow e:/e$   
/diiy + aʔ/  $\rightarrow$  [deeyaʔ] 'one in front'  
/?iidl +iin/  $\rightarrow$  [ʔedlen] 'will hunger'  
/hiwiit + Ø/  $\rightarrow$  [hiwet] 'journey'

$B \rightarrow C$ :  $e:/e \rightarrow a:/a$   
N/A

**Description:** Yowlumne has a complex system of vowel lowering and close syllable shortening, which has been the subject of intense theoretical discussion. Kenstowicz and Kisseberth postulate that Yowlumne has a long vowel system of /i: a: o: u:/, and that long [e:] is always a lowered reflex of /i:/. This generalization has held through to more modern approaches (e.g., D'Arcy (2003)) This would render part of the chain shift analysis postulated in Moreton (2004a) incoherent, as underlying /e:/ does not exist. A more concrete problem is that in the data provided in Kenstowicz & Kisseberth (1979), Archangeli (1991) and McCarthy (1999), there is no evidence of underlying short /e/ lowering to [a]. Surface [a] only appears to be a possible reflex of underlying /a/. Chained lowering effects appear typologically rare, and it would be intriguing indeed if Yowlumne exhibited this process, but it appears that this is not the case.

**Language:** Yokuts (Yowlumne dialect) 2

**ISO Code:** yok

**Language Family:** Yokutsan

**Spoken in:** San Joaquin Valley, California, USA

**Process(es) Involved:**  $A \rightarrow B$ : Vowel Lowering  
 $B \rightarrow C$ : Rounding Harmony

**Domain:** Overlapping

**Schematic:** u:-a  $\rightarrow$  o:-a  $\rightarrow$  o:-o (M)

**Environment(s):**  $A \rightarrow B$ : /\_\_C<sub>(0)</sub>[+low],  $B \rightarrow C$ : /[+round]C<sub>(0)</sub>\_\_

**Source(s):** McCarthy (1999), Moreton & Smolensky (2002), Moreton (2004a)

**Examples:** *examples from McCarthy 1999, p.27*

$A \rightarrow B$ : u:-a  $\rightarrow$  o:-a  
/c'u:m-al/  $\rightarrow$  [c'o:mal] 'might destroy'

$B \rightarrow C$ : o:-a  $\rightarrow$  o:-o  
/do:s-al/  $\rightarrow$  [do:sol] 'might report'

**Description:** Two general effects interact in Yowlumne to form this putative shift. The first is a general rule of long-vowel lowering, which is productive across the language and takes place across with both front and back high vowels. The second is a rule of rounding assimilation, in which “[s]uffix vowels are rounded after a round vowel of the same height” (McCarthy 1999, p.27). That this is a general process can be seen from examples like /dub-mi/  $\rightarrow$  [dubmu] ‘having led by the hand’ (ibid), where it is the high round vowel /u/ that triggers rounding as opposed to mid-round /o/.

#### **4.4 Trends in the corpus.**

Moreton's corpus presents putative shifts in isolation, with no overall assessment of which kinds of shifts tend to recur. This is not especially surprising, as the point of Moreton & Smolensky's article is not to provide a definition of chain shifting.

However, if we *do* wish to define the term, it is vital to know whether there are certain kinds of shift that appear time and again, and whether these more common shifts represent genuine crosslinguistic tendencies.

##### *4.4.1 Language family*

No explicit effort has been made, in the design of the corpus, to have a balance of effects from various language families. The consequence of this is that the corpus reflects the tendency in the Linguistics literature in general to have more readily available data from Indo-European languages than any other family. There are 25 separate language families represented in the corpus (there are 24 families listed, but the label 'Isolate' is attached to two languages which are, obviously, unrelated).

However, I have no reason not to believe that the seeming bias towards Indo-European languages is just that; a bias more indicative of sociological than linguistic factors. I note that in the tables below the numbers refer to putative chain shifts cited in the corpus, not languages. For example, 3 of the 5 Afro-Asiatic shifts listed come from the same dialect, Bedouin Hijazi Arabic.

Table 4-1: Language families in the corpus

Family	<i>n</i>	%
Afro-Asiatic	5	7.58
Algic	1	1.52
Australian	1	1.52
Austronesian	3	4.55
Chukotko-Kamchatkan	1	1.52
Eskimo-Aleut	1	1.52
Hmong-Mien	5	7.58
Indo-European	19	28.79
Iroquoian	1	1.52
Isolate	2	3.03
Japonic	1	1.52
Mayan	1	1.52
Niger-Congo	7	10.61
Nilo-Saharan	1	1.52
Pomoan	1	1.52
Sino-Tibetan	1	1.52
Siouan-Catawban	1	1.52
Tarascan	1	1.52
Totonacan	1	1.52
Uralic	2	3.03
Uto-Aztecan	3	4.55
Wintuan	2	3.03
Yaguan	1	1.52
Yokutsan	4	6.06

Each of the family designations is taken from the highest level of family information available on ethnologue. A slightly cleaner table is presented in table 4-2, in which all of the families from which only one shift is represented are collected together, giving a clearer indication of particularly significant families.

Table 4-2: A cleaned table of language families in the corpus

Family	<i>n</i>	%
Afro-Asiatic	5	7.58
Austronesian	3	4.55
Hmong-Mien	5	7.58
Indo-European	19	28.79
Niger-Congo	7	10.61
<b>Singletons</b>	16	24.24
Uralic	2	3.03
Uto-Aztecan	3	4.55
Wintuan	2	3.03
Yokutsan	4	6.06

Overall, the spread of language families represented is wide, but the amount of shifts in each family is typically low. The only exception to this is Indo-European, which makes up nearly a third of the sample (28.79%).

#### 4.4.2 Domain

Recall that the notion of ‘domain’ is defined over four potential categories: isomorphic, in which the environments for  $A \rightarrow B$  and  $B \rightarrow C$  are exactly the same; adjacent, in which the segments affected by the  $A \rightarrow B$  and  $B \rightarrow C$  processes are next to one another; overlapping, in which no adjacency relation holds, but some portion of the environment of the  $A \rightarrow B$  process is shared with the environment of the  $B \rightarrow C$  process; and non-overlapping, where no obvious relation obtains between the  $A \rightarrow B$  and  $B \rightarrow C$  environments. Table 4-3 shows the relative frequencies of each type of domain in the sample:

Table 4-3: Domains of chain shifting

Domain	n	%
Adjacent	9	13.64
Isomorphic	35	53.03
N/A <sup>21</sup>	4	6.06
Non-overlapping	8	12.12
Overlapping	10	15.15

It is clear that the most common domain is ‘isomorphic’. In fact, more than half of the shifts in the corpus work solely on changing the features of individual segments. The other three domains are all represented nearly equally, accounting for between 10 and 15% of the sample each. At this point, I make no further comment on the nature of the domain of chain shifting, but it is an important topic in the following chapter, where I propose that even those theorists who do not share my position that genuine synchronic chain shifts do not exist would benefit greatly from limiting the domain of chain shifting to the segment.

---

<sup>21</sup> In the following chapter, I will discuss a decision made by Moreton to include in his corpus shifts which are implied on the basis of the OT principle of Richness of the Base. Because they are only implied, there is no data to show what the  $B \rightarrow C$  mapping would constitute in any of these cases. The notion of domain, as it pertains to the sample, is only coherent if we know exactly what both mappings are, and for this reason these putative shifts do not fit into any of the listed domain categories.

#### 4.4.3 Processes involved

Perhaps the most important question in attempting to define chain shift satisfactorily is whether there is some kind of ‘canonical’ chain shift. That is to say, whether there is a particular kind of shift that represents the ‘norm’ for chain shifting, against which other shifts can be measured. There are several metrics by which we may consider instances of a particular phenomenon to be canonical, and as Corbett points out, frequency may not be an especially good predictor of how canonical a certain process is. Indeed, “[t]hey may be extremely rare, or non-existent” (2005, p. 26). Whilst there currently is not, to my knowledge, an explicit definition of how a canonical chain shift might look, by building explicit theories of chain shifting, previous authors have already constructed the “theoretical spaces of possibilities” advocated by Corbett as the first step necessary in taking a canonical approach (2005, p.26). Each chain shift theory is quite explicit in the kind of chain shifts it would allow and not allow in its hypothesis space (see chapter 2 for discussion of this point). It therefore becomes useful to see how well the shifts that we do observe accord with the kinds of shift that an individual theorist might predict, depending on their theoretical stance. Therefore, it is necessary to see which kinds of processes seem to genuinely be involved in chain shifting.

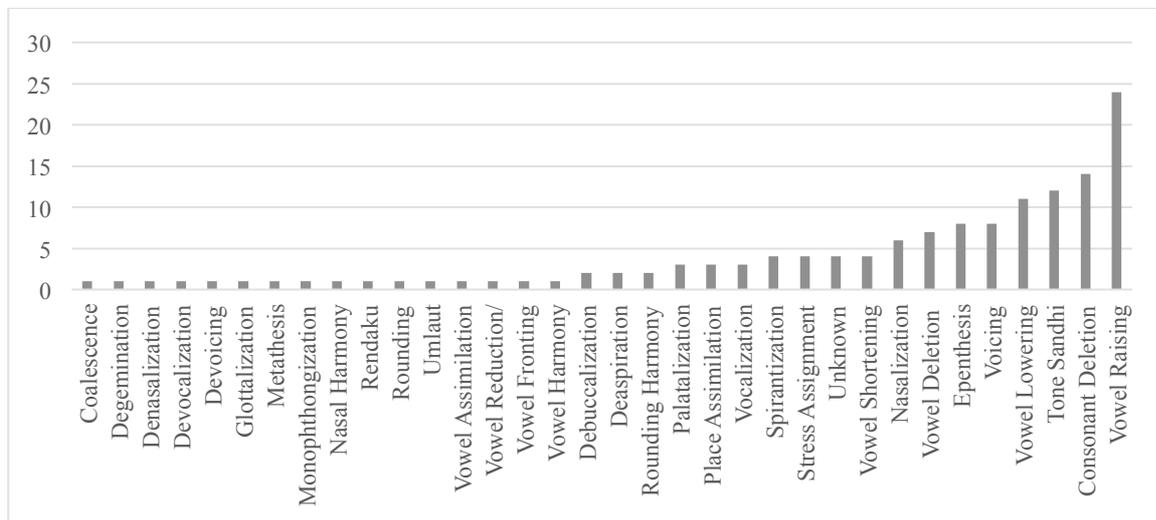
There are 137 phonological processes to be considered overall. There are three shifts in the corpus with more than two parts (Xiamen tone sandhi, Nzebi vowel raising, and Thok Reel vowel lowering) but given that all three shifts are argued to represent unified processes, the extra instances of tone sandhi and vowel raising were not added to overall counts of instances of each process. However, when shifts have multiple possible realizations of the  $B \rightarrow C$  part, like the Finnish shift discussed by Łubowicz (2003a, 2012), in which rounding, vowel lowering and deletion are all possible consequences of the second part of the shift, all of these are listed. Table 4-4, then, is simply the raw count of each individual process by type, regardless of whether that process formed the  $A \rightarrow B$  or  $B \rightarrow C$  part of the shift.

Table 4-4: Raw count of processes in the table

Process	n	%
Coalescence	1	0.73
Degemination	1	0.73
Denasalization	1	0.73
Devocalization	1	0.73
Devoicing	1	0.73
Glottalization	1	0.73
Metathesis	1	0.73
Monophthongization	1	0.73
Nasal Harmony	1	0.73
Rendaku	1	0.73
Rounding	1	0.73
Umlaut	1	0.73
Vowel Assimilation	1	0.73
Vowel Reduction/Deletion	1	0.73
Vowel Fronting	1	0.73
Vowel Harmony	1	0.73
Debuccalization	2	1.46
Deaspiration	2	1.46
Rounding Harmony	2	1.46
Palatalization	3	2.19
Place Assimilation	3	2.19
Vocalization	3	2.19
Spirantization	4	2.92
Stress Assignment	4	2.92
Unknown	4	2.92
Vowel Shortening	4	2.92
Nasalization	6	4.38
Vowel Deletion	7	5.11
Epenthesis	8	5.84
Voicing	8	5.84
Vowel Lowering	11	8.03
Tone Sandhi	12	8.76
Consonant Deletion	14	10.22
Vowel Raising	24	17.52

What is clear from this table is that a great many phonological processes may make up the constituent parts of chain shifts. However, the majority of these processes do not feature regularly in such patterns. Indeed, 16 of the 34 processes listed above occur in only one chain shift process. It seems clear that shifts based on shifts in vowel height (be that raising or lowering), segmental deletion (be that vowels or consonants), and tone sandhi are the most common kinds of shift. Figure 4-2 illustrates both the breadth of processes that can form part of a chain shift and the fact that there are certain processes that are far more commonly found as chain shift constituents.

Figure 4-2: Processes involved in the corpus

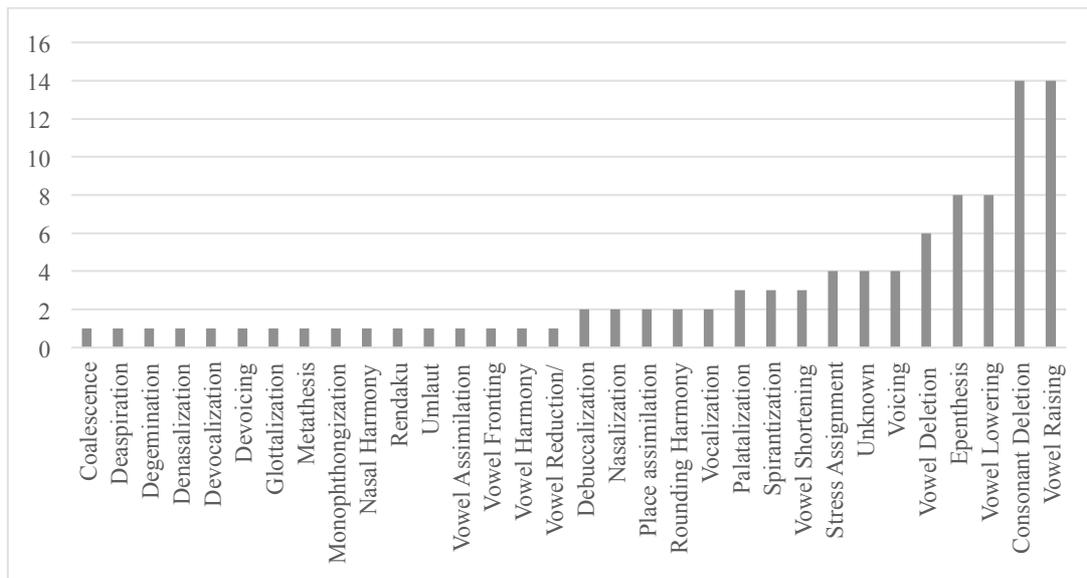


However, it is not sufficient to simply list all of the processes that occur. First, it must be pointed out that, whilst Moreton’s corpus is not constructed with any particular skewing towards any one kind of pattern, the sources that I have used to supplement the corpus have tended to be surveys of particular kinds of effects. Parkinson (1996) and Mortensen (2006) in particular, are concerned only with vowel height and tone sandhi, respectively. Indeed, all of my examples of tone sandhi come from Mortensen, and none were included in Moreton’s original corpus. The charts below show the overall process information taken from only the 47 shifts from Moreton’s corpus that form the bedrock of my sample.

Table 4-5: Processes from Moreton's corpus

Process	n	%
Coalescence	1	1.03
Deaspiration	1	1.03
Degemination	1	1.03
Denasalization	1	1.03
Devocalization	1	1.03
Devoicing	1	1.03
Glottalization	1	1.03
Metathesis	1	1.03
Monophthongization	1	1.03
Nasal Harmony	1	1.03
Rendaku	1	1.03
Umlaut	1	1.03
Vowel Assimilation	1	1.03
Vowel Fronting	1	1.03
Vowel Harmony	1	1.03
Vowel Reduction/Deletion	1	1.03
Debuccalization	2	2.06
Nasalization	2	2.06
Place assimilation	2	2.06
Rounding Harmony	2	2.06
Vocalization	2	2.06
Palatalization	3	3.09
Spirantization	3	3.09
Vowel Shortening	3	3.09
Stress Assignment	4	4.12
Unknown	4	4.12
Voicing	4	4.12
Vowel Deletion	6	6.19
Epenthesis	8	8.25
Vowel Lowering	8	8.25
Consonant Deletion	14	14.43
Vowel Raising	14	14.43

Figure 4-3: Processes involved in Moreton's corpus



The table and graph show that, tone sandhi aside, the overall trend is similar. There are a great many processes that appear once, or very infrequently, whilst effects involving vowel height and segmental deletion are the most frequently recurring processes. The effect of Parkinson's sample is certainly worth noting however, as an additional ten instances of vowel raising have been adduced from it, especially given that in Moreton's corpus, there are exactly the same amount of consonant deletion and vowel raising effects.

#### 4.4.4 Shifts as a whole

Of course, simply showing the raw numbers of all of the processes involved in the shifts in my sample gives an incomplete picture of what is happening in the shifts overall. A first important question to ask is the relative proportion of shifts that consist of two applications of the *same* process, as opposed to two different processes. In order to do this, however, it is important to be explicit about what is meant by the term 'process'. In the exposition so far, 'process' has been a purely descriptive term, with no thought given to the potential motivation behind the changes to the segment or segments that are being acted upon. If this is the definition that we use, then a clear majority of the shifts that we observe (43/66, or 65.15%) are made up of two different processes.

However, when taken together it seems clear that there are certain processes which are not exactly the same operation in terms of, for example, the features that are changed, or the descriptive term that is usually used to describe them, but are examples of the same overall phenomenon. Take, for example, the shift listed above from Gran Canaria Spanish, which takes the form  $p,t,k \rightarrow b,d,g \rightarrow \beta,\delta,\gamma$ . In simple terms, the  $A \rightarrow B$  shift is one of voicing (featurally  $[-\text{voice}] \rightarrow [+ \text{voice}]$ ), whilst the  $B \rightarrow C$  shift is one of spirantization ( $[+\text{voice}, -\text{cont}] \rightarrow [+ \text{cont}]$ ). Different features are changed, thus on the surface the processes are different. However, this seems to ignore a fairly clear generalization; taken together, the effect in Gran Canaria Spanish appears to be one of lenition. Lenition has often been characterized as a single process with multiple, related reflexes. In some approaches to phonology, most notably Element Theory (see e.g. J. Harris & Lindsey (1995) or Backley (2011) for an introduction to the theory), most reflexes of lenition are expressed in exactly the same way; the reduction of segmental complexity via the deletion of elements (see, e.g., J. Harris (1990)). Taking influence from this approach, there are 11 processes in the corpus that can best be described as examples of Lenition.

Additionally, there are 15 consonantal processes that do not count as unified lenition trajectories, but rather seem to be totally assimilatory in nature. Usually, this takes the form of consonant voicing (a typical example of lenition) followed by nasalization (see, for example, Manyá or Nzema). Finally, there are four potential examples of effects that could constitute two unified processes of vowel reduction, in Bedouin Hijazi Arabic and Palauan, even though the two processes involved are not the same.

Crucially, this relabeling of certain processes has consequences for the count of shifts that are formed of two instances of the same process. Under these definitions, 53.03% of shifts in the corpus (35/66) are reapplications of one process, whilst the other 46.97% remain effects made up of two separate operations.

Of the shifts that are unified, in the sense that they are made up of the repeated application of the same process, it is clear that some trends are more prevalent than others:

Table 4-6: Unified processes in the corpus

Process	<i>n</i>	%
Consonant Deletion	1	2.86
Vowel Reduction	2	5.71
Vowel Deletion/Vowel Shortening	3	8.57
Vowel Lowering	3	8.57
Assimilation	4	11.43
Lenition	5	14.29
Tone Sandhi	6	17.14
Vowel Raising	9	25.71

Vowel raising, tone sandhi, and lenition processes are the three most prevalent kinds of shift in the sample. Indeed, if vowel lowering is conflated with vowel raising, given that both kinds of processes, at least on the surface, both simply represent minimally different kinds of vowel shift, then the overall category of vowel shift accounts for more than one third of the examples in the corpus (12/35, 34.29%). Strikingly, whilst segmental deletion processes are, taken together, the second most common kind of chain shift process (21/136, 15.44%), they form unified processes of shifting only 4 times (11.42%).

With regard to finding a canonical definition of chain shifting, it is clear that it is not enough simply to say that shifts in vowel height are canonical simply because they are the most common kinds of shift in the corpus. However, as shown in Chapter 2, vowel raising shifts can be modeled by all major chain shift theories, and are frequently used as an example of the phenomenon as a whole (see, e.g., Kirchner (1996), Gnanadesikan (1997), Moreton & Smolensky (2002), Łubowicz (2003a), Mortensen (2006) etc.). For this reason, coupled with their high frequency in both Moreton's compendium and my new, wider sample, I will be returning to vowel height shifts at various points throughout this thesis.

#### 4.5 Where do we go from here?

Throughout the new corpus I have presented the putative shifts without much in the way of comment on whether I believe they constitute chain shifts. This is primarily because, as previously stated, I do not believe that synchronic chain shifts exist.

Why, then, have I spent a chapter giving an extensive listing of all of the putative examples of synchronic chain shift that I can find? There are two main reasons. The first is that I do not assume that this thesis will be the last word in chain shift studies. Because of this, I believe that updating Moreton's compendium is a worthy exercise in its own right, as it gives anyone interested in researching synchronic chain shifts a newer, expanded collection of chain shift specimens to analyze under whatever theoretical analysis they see fit.

The second is that, in order to argue against the concept of synchronic chain shifting, it is necessary to have as full an understanding as possible of what is being argued against. It is all very well to assert that there is no such thing as a synchronic chain shift, but unless the arguments that I marshal against the concept hold good no matter what kind of chain shift is under discussion, then the arguments are far less valid. I believe that this is a major problem with previous proposals of how to deal with chain shifting, as they tend to either focus heavily on one particularly interesting chain shift (see Łubowicz (2003a), whose chain shift theory is largely based on her extensive discussion of Finnish), or one specific class of processes (a good example being Mortensen (2006), which almost exclusively discusses tone shifts). Even if these theories can model a wide variety of shifts, the fact that they do not attempt to address the diversity of putative chain shifts that have been claimed to exist in the literature makes a strong prediction that those unaddressed, unstudied, shifts have the same motivation and *should* be modeled in the same way as those shifts that are considered more fully.

Since Neasom (2011), I have been skeptical of the notion that such a wide-ranging set of processes benefits from being lumped together under the banner of synchronic chain shifts. In the following chapter, I argue that processes that have been called chain shifts have either been misdescribed or can be more accurately represented as epiphenomena of other processes.



## Chapter 5: Reasons to doubt the existence of synchronic chain shift

### 5.1 Introduction

In the previous chapter I presented a sample of putative synchronic chain shifts. To my knowledge, this is the largest such sample yet assembled. Because of the ‘big-tent’ approach taken in my construction of the corpus, it represents not only the widest sample of individual processes that currently exists, but certain shifts within it also represent the widest *definition* of chain shifting that currently exists. The overall aim of this chapter is to argue that, of the 66 putative shifts in the corpus, many can be completely discounted as potential examples of synchronic chain shift. I attempt to illustrate this through detailed case studies of individual examples of shifts from the corpus, showing that chain shift analyses are inappropriate in these cases. There are many potential reasons that a particular effect might need reanalysis, and these reasons may not even be linguistic.

In the following section of this chapter I discuss shifts that we can dismiss based on practical factors, such as a lack of data or a reassessment of whether the shift in question should even be considered to be synchronic. In this section, I also look at the category of shifts, all from Moreton’s corpus, in which the C part of the shift is schematized as ‘unknown’, and I argue that these too should be discarded. The third part of the chapter is the first of a two-part discussion of chain shift as an epiphenomenon. The first part of this discussion focuses on the domain of chain shifting, arguing that putative chain shift processes above the level of the segment should be discarded from the sample. The second part discusses a different kind of epiphenomenal behaviour; accidental chain shift as a small part of a wider process.

Taken to its logical conclusion, the discussion of this sort of epiphenomenal chain shift casts doubt even on examples like Lena Spanish, which appear to be relatively ‘canonical’ chain shifts (i.e., every observed alternation appears to be in service of the chain shift). Therefore, this chapter concludes that there is no strong argument for calling any of the processes in the sample in the previous chapter a synchronic chain shift. The main aim of this chapter is to make clear that putative examples of

synchronic chain shift are either problems that do not need solving, or problems that we can solve without recourse to specific chain shift methodologies.

## 5.2 Practical problems

Before discussing theoretical problems with particular kinds of shift, it is well worth looking at more practical issues. Recall Moreton's quote about how the compendium of shifts came about; "Some [shifts] are taken from the secondary literature; others were found by flipping through reference grammars in the library looking for telltale phrases like "This rule does not apply to long vowels created by Rule 27"" (2004a, front matter). This first section of the chapter represents a more thoroughgoing 'flip through' of some of the less well-reported shifts in the corpus, and gives evidence against certain of the included shifts being plausible candidates for synchronic shifthood. Where possible, these problematic shifts are taken from the set of vowel raising processes (or its inverse, vowel lowering). As discussed in the previous chapter, vowel height shifts are perhaps the most 'canonical' kind of chain shift that we observe. They are frequent relative to other kinds of shifts, they operate at the level of the segment, and they have a seemingly unified reflex. That even some of the vowel raising examples appear to be problematic illustrates that the issues with the processes in Moreton's corpus 'go all the way to the top'.

### 5.2.1 *A lack of synchrony: Kashubian, African-American Detroit English*

It is trivially true that it is important for any putative synchronic chain shift to be genuinely synchronic. As discussed in chapter 3, there are substantive differences between how shifts in synchrony, diachrony, and acquisition operate. However, some of the entries in the new chain shift corpus do not appear to be genuinely synchronic shifts. I discuss two key examples of this below.

#### 5.2.1.1 *Kashubian*

Łubowicz (2003a) discusses a set of differences between standard Polish and Kashubian, spoken in the Pomerania region of Poland<sup>22</sup>. In Kashubian, words that are realized with surface [a] in Standard Polish are sometimes realized with [e], and

---

<sup>22</sup> Łubowicz describes Kashubian as a Polish dialect, but Hopkins (2001) and Sanders (2003) suggest that it may well be an entirely separate language. Hopkins (2001, p.3) makes the point that this decision depends more on external factors than genuinely linguistic ones.

words that are realized with surface [e] in Standard Polish also raise. The target of this second raising is not consistent, with realizations of [i], [ɪ], [ie], [iy], [é] (described by Łubowicz as a raised /e/), and [i̯] all reported as regional variants (2003a, p.48). Łubowicz gives the examples in (1a-g) (2003a, pp.48-49), based on Urbańczyk (1972), and Dejna (1993):

(1)	STANDARD POLISH	KASHUBIAN	GLOSS
(a)	tr[a]va	tr[e]va	‘grass’
(b)	pt[a]k	pt[e]χ	‘bird’
(c)	g[a]d[a]	g[e]d[e]	‘he talks’
(d)	por[a]nek	por[e]nk	‘morning’
(e)	bz[e]k	bz[ɪ]k	‘shore’
(f)	s[e]r	s[ɪ]r	‘cheese’
(g)	ml[e]ko	ml[ɪ]ko	‘milk’

Łubowicz does not explain why certain vowels are raised in Kashubian and why others are not. For example, in (1a), tr[a]va → tr[e]va, the second vowel is untouched by the putative raising process. But it is not the case that only one vowel per word can raise, as (1c), g[a]d[a] → g[e]d[e], shows. (1c) also shows that there is no proscription against raising of final vowels. This raising effect does not appear to be systematic, in the way that one would expect if this were an active phonological process. The examples in (1a-g) as a whole do not give any indication of a unified trigger for the raising. It is hard to understand how a speaker would learn a productive rule governing these alternations.

A more general concern is that Łubowicz, makes the strong theoretical claim that Kashubian and Standard Polish share underlying representations. She justifies her decision thus; “[t]he evidence for this comes from borrowings, some of which surface with Standard Polish vowels (Lorentz, 1958)” (p. 47). It is true that Kashubian speakers are almost exclusively bilingual, indeed Topolińska finds only one monolingual Kashubian speaker in her study of the language (1974). However, in the experimental literature on bilingualism “a consensus has emerged that languages have functionally separate stores for form-based, lexical (phonological and orthographic) representations” (McElree, Jia, & Litvak, 2000, p. 229). This

suggests that there are separate word stocks in the minds of Kashubian speakers for Kashubian and Polish words. It is simply the case that, when a speaker is speaking Kashubian, they use the Kashubian form.

It is also worth bearing in mind that there are “significant differences” between Kashubian and Standard Polish (Hopkins (2001, p. 3)), that make Łubowicz’s schematic hard to interpret even if we follow her assumption that Kashubian forms are derived from Standard Polish ones. Figure 5-1 and figure 5-2 show the vowel systems of the two languages:

*Figure 5-1: Kashubian vowels (Hopkins, 2001, p. 23): /i é e ə á ó o u/*

i            u  
 é    ə    ó  
 e    á    o  
 a

*Figure 5-2: Standard Polish vowels (Łubowicz (2003a, p.47) (citing Rubach (1984)): /i ɨ e a o u/*

i    ɨ    u  
 e    o  
 a

There are striking differences between the systems in Kashubian and Standard Polish. For instance, Kashubian has [+/-ATR] variants of its low vowel (/a á/), and its mid-vowels (/e é o ó/), a distinction not present in standard Polish. As the relationship between the vowel systems of standard Polish and Kashubian is not isomorphic, and this is not acknowledged by Łubowicz, it is difficult to know exactly what her mappings in (1a-g) are supposed to represent. In sum, I dispute the notion that there is any Kashubian speaker with a coherent /a/ → [e] → [ɨ] chain shift as part of their synchronic grammar.

### 5.2.1.2 African-American Detroit English

Another example of a putative chain shift that cannot reasonably be said to be synchronic comes from African-American Detroit English (AADE). Moreton lists a shift with the schematic form  $\text{aid} \rightarrow \text{a}^{\text{e}}\text{d} \rightarrow \text{ait}$  in his compendium, and references a conference presentation by Bridget Anderson (1999). Anderson also published a paper on the same topic in the *Journal of Sociolinguistics* in (2002), on which I base my discussion. The key finding of this paper is that younger AADE speakers reliably monophthongize /ai/ before voiceless obstruents, which is not true either of older speakers of AADE, or speakers of other varieties of African-American English. In all AAE dialects, monophthongization before voiced obstruents is common (/ta:d/ for *tied*), but it is only in the AADE dialect that this has spread to pre-voiceless contexts (/ta:t/ for *tight*). Anderson postulates that this is because of dialect contact between AADE speakers and white Southern Appalachian speakers, who are the main white community in central Detroit, and who also monophthongize reliably before voiceless obstruents.

It is difficult to see where the inspiration for Moreton's  $\text{aid} \rightarrow \text{a}^{\text{e}}\text{d} \rightarrow \text{ait}$  schematic comes from. The  $\text{A} \rightarrow \text{B}$  part of the shift seems to be incoherent in that, for all groups of speakers (Anderson collected data from 27 participants, aged between 4 and 81), the dominant pronunciation in pre-voiced contexts was [a:] or [a<sup>e</sup>]. Monophthongization in this context is the dominant pronunciation not just in AADE speakers, but all AAE speakers, and white Southern speakers too. Therefore, it seems more likely that the underlying form for a word like *tied* would, in fact be /ta:d/ or /ta<sup>e</sup>d/. This is supported by further data from Anderson's study, in which the dominant pronunciation for the /ai/ diphthong in all contexts for AADE speakers is the monophthongized version, including before sonorant consonants and pre-pausally (2002, p. 92). The first part of the shift, then, seems more likely to be illustrating the trajectory of a change in progress, rather than a synchronic alternation.

Even if one wished to maintain that all instances of the monophthong [a:] were based on underlying /ai/, and that this was a synchronic alternation shared by all speakers, the  $\text{B} \rightarrow \text{C}$  ( $\text{a}^{\text{e}}\text{d} \rightarrow \text{ait}$ ) part of the shift is still troubling. The first issue is that Anderson does not discuss any consonant voicing alternations. Therefore, it is

difficult to see what kinds of alternations Moreton has in mind. Secondly, the mapping shown would seem to go against the argument, made by Anderson, that monophthongization is spreading. In the schematic, Moreton appears to be suggesting that underlying sequences of monophthongal or glide-reduced /a:/ and an underlyingly voiced obstruent undergo two changes; diphthongization and devoicing. These changes are not even hinted at in Anderson's article. A third, related point is that Moreton's schematic features a glide-reduced /a<sup>ɛ</sup>d/ as its B part. It is not clear that this is a potential underlying structure in any dialect of English, and Anderson is unconcerned with this issue in her article. The overriding point is that, even if there are alternations that Moreton is aware of that Anderson does not discuss, this is still a shift that bears all of the hallmarks of a shift in progress as opposed to a genuine synchronic shift.

### 5.2.2 Practical issues: *Insufficient or unclear data*

As I have previously discussed (see chapter 4), Moreton's compendium of shifts does not include examples of the parts of each individual shift. In finding examples to illustrate the shifts in my own sample, it has become clear that in certain cases, in both Moreton's compendium and other sources, the data is very scant, or very unclear. Putative shifts based on troublesome data may well give the impression that the set of synchronic chain shifts is wider than it actually is, or that certain processes can take part in chain shifts when there is no substantive evidence that they do. This section discusses two further examples where the data gives insufficient evidence that there is any genuine A → B → C order.

#### 5.2.2.1 Basque (*Etxarri Navarrese and otherwise*)

Moreton & Smolensky (2002) use a putative a → e → i shift in Western Basque to exemplify the mechanics of a chain shift. They take their data from Kawahara (2002, p.78), which is represented below:

(2)	INDEFINITE	DEFINITE	GLOSS
(a)	alaba bat	alabea	'daughter'
(b)	neska bat	neskea	'girl'
(c)	seme bat	semie	'son'
(d)	ate bat	atie	'door'

Kawahara cites Kirchner's 1996 article in *Linguistic Inquiry* as the source of this (rather sparse) data. However, Kirchner (1996) uses data from a vowel raising shift in Nzɛbi (see chapter 4) as his example of a raising shift, and does not list the Western Basque examples (though he does mention "Basque vowel raising under hiatus", referencing Hualde (1991) (1996, p.344)). Moreton & Smolensky also reference an earlier draft of Kirchner's work on synchronic chain shifts (1995), which does make more explicit reference to Western Basque. This discussion makes clear that the data is not as simple as it appears in Moreton & Smolensky (2002) and Kawahara (2002). Firstly, the full data set shows that there are three additional mappings;  $o \rightarrow u$ ,  $i \rightarrow i^y$ ,  $u \rightarrow u^w$  (data below from Kirchner (1995, p. 5)):

(3)	(a)	asto bat	astue	'donkey'
	(b)	basto bat	bastue	no gloss given
	(c)	erri bet	erri <sup>y</sup> e	'village'
	(d)	ari bet	ari <sup>y</sup> e	'thread'
	(e)	buru bet	buru <sup>w</sup> e	'head'
	(f)	iku bet	iku <sup>w</sup> e	'fig'

This would suggest that, as in Nzɛbi, the effect in Western Basque would constitute a three step shift, with low vowels becoming mid, mid vowels becoming high, and high vowels raising and diphthongizing. However, Kirchner makes clear that this may not be the case. The datasets above conflate several Western Basque dialects, and it is not certain that a coherent  $a \rightarrow e \rightarrow i \rightarrow i^y$  shift is present in any one dialect. The clearest apparent shift is found in the Etxarri Navarrese dialect, which does not include the  $/a/ \rightarrow [e]$  mapping. In this dialect,  $/a/$  surfaces as  $[a]$  in the hiatus context where mid and high vowels undergo raising. It is important, therefore, to point out that whilst this does not preclude a chain shift analysis of Etxarri Navarrese Basque, it does mean that the accounts given in Kawahara (2002) and Moreton & Smolensky (2002), are problematic in two ways. First, they ignore the final ( $i \rightarrow i^y$ ) stage of the shift. More importantly, they suggest that there is at least one dialect in which  $/a/ \rightarrow [e]$  and  $/e/ \rightarrow [i]$ . It is not certain that this is the case.

The final diphthongization stage is unique in the vowel raising mappings in this sample, as it is essentially a process of diphthongization, rather than genuine raising. Whilst diphthongization is a common part of diachronic shifts (see, e.g., the discussion and references on the GEVS in chapter 3), there are no other putative synchronic shifts that feature this kind of alternation. More generally, it is also the only vowel raising shift (*pace* Kashubian Polish, see previous section), in which one of the mappings is not structure preserving. Usually the highest phonemic vowel of the system is the termination point for the vowel raising effects, as in every other effect in the sample. Kirchner’s OT analysis does not suggest a reason for this, and indeed does not address the question. Whilst his analysis makes clear that raising takes precedence over any other concern, it is not clear that diphthongization achieves this aim in a genuinely phonological sense. Whilst, given Kirchner’s functionalist approach to OT, it makes sense within his framework to suggest that the phonology might be able to attend to the difference in height caused by the co-articulation of, say, /i/ and the glide /y/, this requires a theoretical move that is highly unusual. Basque is usually envisioned as a five-vowel system, at least in most dialects, including Etxarri Navarrese Basque. It has the standard set of five vowels /i e a o u/ (see, for example, Hualde 1991 p.11).

Hualde states that “[i]n general, there are no underlying glides” in Basque (p.11), and that when glides do appear on the surface, they are simply realizations of the high vowels /i/ and /u/ appearing in certain positions (post-vocalic in all dialects, and in various other positions depending on dialect). This suggests that these glides should be specified in the same way as the vowels, only affiliated to consonantal positions rather than vocalic ones. Kirchner departs from this analysis by suggesting that Basque vowels are possessed of a feature [+/- raised], and that the diphthongs [i<sup>y</sup>] and [u<sup>w</sup>] are the only [+ raised] vowels in the system. Table 5-1 is from Kirchner (1995), and shows his conception of the vowel system:

*Table 5-1: Kirchner’s model of Etxarri Navarrese Basque vowel height (1995, p.6)*

	Low	High	Raised
i <sup>y</sup> , u <sup>w</sup>	-	+	+
i u	-	+	-
e o	-	-	-
a	+	-	-

As far as I am aware, this [+/- raised] feature is of Kirchner's own design. He does not give a source for the feature, and no other accounts of Basque that I have read include it (certainly not Hualde (1991), de Rijk (1970), or Kenstowicz and Kisseberth (1979), who are the sources that Kirchner cites). Given Hualde's position that glides in Basque are not underlying, and given that the *i<sup>y</sup>* and *u<sup>w</sup>* diphthongs are, in all probability, not part of the phonemic inventory of Basque, invoking an entire new phonological feature purely to deal with such diphthongs is not a plausible way to explain the differences between these sounds and the standard high vowels [i] and [u].

In terms of what Etxarri Navarrese Basque can tell us about synchronic chain shift, the question is whether the raising observed in the high vowels during diphthongization is phonological, or whether it is just an epiphenomenal phonetic effect of the diphthongization. Kirchner's analysis crucially relies on the idea that the raising will be accessible to the grammar. As in his analysis of Nzebi, Kirchner uses a RAISING constraint (in this case a more specific HIATUS RAISING constraint), which gradiently assigns violations based on how far from the highest vowel in the system the vowel in a given surface realization is. In Etxarri Navarrese Basque, [i<sup>y</sup>] is adjudged to be the highest vowel in the system, incurring no violations of HIATUS RAISING, with [i] incurring one violation, [e] incurring two, and [a] incurring three. It feels like a dangerous move to allow levels of height that are only present on the surface in predictable environments to affect how constraints can manipulate input forms.

It is also problematic that Kirchner's analysis does not deal with the diphthongization that changes /i/ to [i<sup>y</sup>]. No reason is given why the HIATUS RAISING constraint does not simply force a higher, but monophthongal, version of /i/, as opposed to causing diphthongization as well. As there are issues with both the /a/ → [e] and /i/ → [i<sup>y</sup>] part of the shift, vowel raising in Western Basque cannot be said to be a reliable example of a synchronic chain shift, at least not one in which both A → B and B → C mappings unambiguously involve phonological raising effects.

#### 5.2.2.2 *Pero*

*Pero*, a Chadic language spoken in Nigeria, exhibits a complex system of vowel height effects, which is discussed in detail in Frajzyngier's grammar of the language (1989). All but one of these effects are one-step assimilatory processes. However, Frajzyngier postulates the existence of a two-step lowering shift in a specific phonological context. Frajzyngier's rule for this process (1989, p. 40) is given below in (4):

$$(4) \quad V \rightarrow [-1 \text{ Height}] / V(C)\_ \text{CCV}$$

This rule states that a vowel lowers by one step after a vowel (with an optional following consonant) and before a CCV cluster. Examples of this process are given below:

- (5) (a) /ní ðll + kò/ → [niéllòyò] 'I stood up'  
(b) /cì mà + yì + n + nò/ → [cimàyénnò] 'if you (f) don't make for me'  
(c) /àn + céngèw/ → [ánjáŋgèw] 'one who is stubborn'  
(d) /nì + céyy + kò/ → [nìjáyóyò] 'I drunk all of it' (1989, p. 40)

However, Frajzyngier's *n*-ary rule in (4) is problematic in and of itself (see chapter 2.3.1 for more detailed discussion of this point). As well as this, the environment for the rule is both heavily restricted and essentially arbitrary. Frajzyngier does not give a reason why this environment should cause the lowering effect. Additionally, even within the 'Phonology' chapter of the grammar, it is possible to find counterexamples for both the A → B (high → mid) and the B → C (mid → low) parts of the shift.

- (6) (a) /àn/ + /úgto/ → [ànúgrò] 'one who uproots' (p.12)  
(b) /àn/ + /íppò/ + /díò/ → [ànífòdìò] 'bird catcher' (p.12)  
(c) /cirép/ + /mu/ → [círémmu] 'our women' (p.23)

With such small samples of data on both sides, it is difficult to make further comment. However, indirect evidence for the position that high and mid vowels can

surface faithfully in the proposed lowering environment comes from a Pero-English wordlist, also compiled by Frajzyngier (1985). Examples (7a-7k) below illustrate that the V(C)\_CCV structure in which the rule is supposed to apply does not always condition the lowering that one would expect if the rule is intended to be productive.

- (7) (a) àrítumbà ‘blade of a hoe’ (p.20)  
 (b) òàlínýé ‘tiredness’ (p.21)  
 (c) jérindé ‘hornet’ (p.33)  
 (d) kíjímò ‘like this’ (p.36)  
 (e) kodínye ‘refusal’ (p.36)  
 (f) kúrùm cílìN ‘bladder’ (p.39)  
 (g) mùlínýàn ‘member of the same clan’ (p.42)  
 (h) pàringà ‘deity, believed to cause death and make life’ (p.44)  
 (i) pilya múrbì ‘place name in Filiya’ (p.46)  
 (j) èlénkò ‘because’ (p.30)  
 (k) pendéngrà scorpion (p.45)

Whilst the appearance of word-forms such as those in (7a-k) is not definitive evidence against the formulation of Frajzyngier’s rule, there does not appear to be a pressure in the language against the structure that the rule prohibits. As there is very little information provided about the effect that is not directly contradicted by other examples from the data, my position is that it is unwise to use a case like Pero as a piece of evidence for or against any particular theory of chain shifting.

In this section, I have shown that there are several ways in which seeming examples of synchronic chain shifts are not as clear cut as they would appear simply from looking at their shift schematic, or even a handful of examples. In the next subsection I address a different sort of problem; chain shifts that are only admissible as examples if certain, theory-specific considerations are taken into account.

### *5.2.3 Leaps into the unknown: Chain shifts and Richness of the Base*

Throughout this thesis, I have attempted to be relatively theory neutral. If we are to find a set of chain shifts that are actually worthy of analysis, the data should be, at least at first blush, amenable to explanation using any derivational theory. Note that

this does not mean that all such theories should be able to model the shifts equally *well*. Indeed, the hope would be that there will be a particular theory of chain shifting that is simpler, more parsimonious, and more complete than any other. For genuine theory comparison to take place, though, it is important that the data set under discussion should be the same for all of the theories that we are comparing.

This last statement encapsulates an important concept in comparing theories of chain shifting. This is largely because of a decision made by Moreton in his corpus to include putative shifts in which the C part of the shift must be listed as ‘unknown’. The reason for these unknowns is that Moreton includes examples of shifts that must be inferred through the Optimality Theoretic principle of Richness of the Base (Prince & Smolensky, 1993/2008). The argument for the inclusion of shifts of this nature runs like this:

(a) Chain shifts are  $A \rightarrow B \rightarrow C$  mappings. In OT terms, this means that some input (A, B, or C) has some corresponding output (B or C).

(b) According to Richness of the Base, if the constraint set is correctly calibrated, then *any* input whatsoever should lead to some harmonic output. This is assumed to hold even for inputs that are not phonemes of the language in question.

(c) This means that in a situation where surface [B] is only ever derived from underlying /A/, we still have to assume that /B/ is a *possible* input.

(d) We have no direct evidence for what underlying /B/ should map to, but given that the only source of surface [B] is from underlying /A/, we can infer that an underlying /B/ does *not* map to a fully faithful version of itself, surface [B].

(e) Given that /A/ surfaces as [B], and our hypothetical underlying /B/ does not map faithfully to surface [B], we are forced to assume that there is a chain shift of some sort at work here. However, there is no way of directly observing the /B/  $\rightarrow$  [C] mapping of that chain shift, hence its designation as ‘unknown’.

I would argue that this prediction is uninteresting, because it is unfalsifiable. There are four examples in Moreton’s corpus that have the schematic form  $A \rightarrow B \rightarrow \text{unknown}$ . I restrict my discussion of these to a coalescence process in Sanskrit, for two reasons. The first is that it is the process that has had the most attention from other theorists (see especially Gnanadesikan (1997) and McCarthy (2005)). The second is that my arguments against the Sanskrit effect being seen as a chain shift can be assumed to hold good for the other examples of shifts where the quality of the C part is unknown, as they are fairly general.

In Sanskrit, there is a process of coalescence whereby /a + i/ sequences and /a + u/ sequences are realized on the surface as [e:] and [o:] respectively. As McCarthy points out (2005), this is a fairly normal kind of coalescence. Length is preserved (two short vowels come together to make one long vowel) and the quality of the coalesced vowel is the midpoint of its two constituent parts (data from Gnanadesikan (1997, p.140)):

- (8) (a) /ca + ihi/  $\rightarrow$  [ceeha] ‘and here’  
 (b) /ca + uktam/  $\rightarrow$  [cooktam] ‘and said’

This, apparently, is the *only* source of [e:] and [o:] in Sanskrit. It is also important to note that both McCarthy and Gnanadesikan report there are no instances of short [e] and [o] at all in the language. If this is true, then from a theory neutral perspective we need know nothing else about the Sanskrit effect. If there are no instances of underlying /e:/ and /o:/, then the effect cannot be a chain shift. This is simply because whilst there is clearly an  $A \rightarrow B$  mapping (/ai/  $\rightarrow$  [e:]), a  $B \rightarrow C$  mapping is impossible as there are no instances of underlying B. However, this is incompatible with Richness of the Base.

On Gnanadesikan’s analysis, the Sanskrit effect very much is a chain shift. She assumes the same Vowel Height scale that she uses for her discussion of Lena Spanish (see chapter 2 for discussion of this), where /a/ has the value LOW, /e/ and /o/ have the value MID, and /i/ and /u/ have the value HIGH. She also uses markedness constraints that work to rule out diphthongs (NO-DIPH) and Mid Vowels (\*MID V). Crucially, NO-DIPH outranks \*MID V. This means that, for an input /a +

i/, [e:] is preferred to fully faithful [ai]<sup>23</sup>. The other potential outputs, [i:] and [a:], are ruled out by the chain-shift constraint IDENT-ADJ(VH), which assigns violations to candidates where the output value for vowel height is not adjacent to the input value on the scale. Gnanadesikan follows the standard approach to correspondence theory in which every input segment must have a correspondent of some kind. Thus, a mapping of /a<sub>1</sub>i<sub>2</sub>/ → [a<sub>1</sub>a<sub>2</sub>] would violate IDENT-ADJ because of the /i<sub>2</sub>/ → [a<sub>2</sub>], HIGH to LOW mapping it contains. Conversely, an /a<sub>1</sub>i<sub>2</sub>/ → [i<sub>1</sub>i<sub>2</sub>] mapping would violate IDENT-ADJ because of the /a<sub>1</sub>/ → [i<sub>1</sub>] LOW to HIGH mapping. The high ranking of the IDENT-ADJ(VH) constraint is key to ensuring that [e:] is chosen as the output form:

Table 5-2: Underlying /ai/ on Gnanadesikan’s analysis (1997, p.142)

/a <sub>1</sub> i <sub>2</sub> /	IDENT-ADJ(VH)	NO-DIPH	*MID V	IDENT(VH)
aa <sub>1,2</sub>	*!			
ii <sub>1,2</sub>	*!			
a <sub>1</sub> i <sub>2</sub>		*!		
☞ ee <sub>1,2</sub>			*	**

Of course, given the principles of Richness of the Base, we also have to attend to what would happen with some imagined underlying /e/. Gnanadesikan presents the tableau in table 5-3.

Table 5-3: Underlying /e/ on Gnanadesikan’s analysis (p.143)

/e/	IDENT-ADJ(VH)	*MID V	IDENT(VH)
☞ a			*
☞ i			*
e		*!	

The point that this tableau makes is that a mid-vowel surfacing faithfully is a worse outcome than a surface form for input /e/ that differs in vowel height. Gnanadesikan does not give a preference for whether [i] or [a] is the preferred surface form, saying

<sup>23</sup> From this point onwards, as in Gnanadesikan and McCarthy’s analysis, I will use only /ai/ and /e:/ as examples, even though the pattern also affects back vowels (/au/, /o:/). This is purely as a space saving measure. All arguments that hold for front vowels hold equally for back vowels.

that “other constraints will determine which one” (p.143). What those constraints are is left to the reader to work out. An argument in favour of a realization of surface [a] could potentially come from the idea that low, sonorous vowels like *a* are unmarked relative to high vowels like *i* and *u* (see, for example, Prince & Smolensky (1993/2008), who instantiate the sonority hierarchy in OT), and in a situation where you would otherwise get a tie, low-ranked constraints enforcing universal markedness relations step in (these kinds of constraints are discussed by, for example, de Lacy (2006)). However, even if we assume this universal hierarchy, it is trivially true that there are other constraints that can outrank the constraints that enforce the hierarchy. Were this not the case, there would be only /a/ vowels in the world’s languages. Without engaging in far deeper analysis of Sanskrit than the scope of this thesis allows, it is not possible to decide whether [a] or [i] should be the preferred output.

McCarthy (2005) discusses Gnanadesikan’s analysis of Sanskrit at length, using it as an exemplification of his principle of free rides. In a free ride, learners are presented with positive evidence that a particular structure comes from a productive alternation. In this case, the generalization is that surface [e:] is predictably created from the interaction of /a/ and /i/ at a morpheme boundary. McCarthy argues that, in the process of phonological acquisition, Sanskrit learning children reanalyze cases of surface [e:] that do not occur at a morpheme boundary, assuming that *all* instances of surface [e:] come from input /a+i/. The key piece of evidence that McCarthy presents for this somewhat abstract solution is that there are no instances of surface short [e] in Sanskrit, meaning that the only positive evidence that speakers ever get for the source of any kind of mid vowel comes from examples like /ca + ihi/ → [ce:hi]. Speakers can see that the two individual morphemes feature short [a] and short [i] respectively, when they are pronounced in isolation. Whenever they are pronounced together, [e:] results. Because short [e] *never* surfaces, there is no positive evidence for any kind of mid vowel *not* being the product of coalescence. Thus, even though it seems a more abstract solution, the free ride account is (according to McCarthy) more easily learnable than an account with underlying /e:/.

Another advantage of McCarthy’s analysis is that we solve the problem of having to make an unfalsifiable prediction about what an underlying /e:/ would map to,

because speakers gradually refine their grammar such that their stable adult representations contain no underlying instances of /e/, whether short or long. My reading of McCarthy, is that Richness of the Base is present in the initial grammar, but that certain underlying forms are categorically ruled out once learning is complete. This means that, under McCarthy's account, the schematic for the Sanskrit effect would be /a + i/ → [e:] and nothing else. As in a rule-based derivational account, there is no need to postulate a map for an underlying /e:/ which cannot exist. There is thus no chain shift here

This raises a further question. Perhaps a chain shift is not present in the final state, but is there a coherent shift present at an intermediate stage of learning? McCarthy provides tableaux for his analysis of Sanskrit, which show each stage of learning. In the initial state, the learner, on hearing surface instances of [e:], can be taken to assume a faithful underlying form /e:/. In the initial state, and subsequent tableaux that illustrate stages of learning before the free ride has taken hold, underlying /e:/ always maps faithfully to [e:]. This suggests that, on McCarthy's analysis at least, there is no stage of phonological learning where an A → B → C mapping can genuinely be observed.

There are three other potential cases of chain shifts assumed on the basis of Richness of the Base in Moreton's corpus: two from dialects of Yokuts (one from the Wikchamni dialect, the other from the Yowlumne dialect), and one from the Algonquian language Ojibwa. On the basis that, as in Sanskrit, there is no direct evidence in any of the three cases that there is a complete chain shift, I believe that I am justified in disregarding these examples from any further analysis that I undertake. In sum, for an effect to be considered a potential chain shift in this thesis, there must be direct evidence of distinct surface forms for *all three* parts of the chain shift, A, B and C.

### 5.3 Domain restriction

In the previous chapter, I noted that certain kinds of process seem to recur in the chain shift literature. Two of these kinds of processes are changes in vowel height and the manner of consonants. This is in line with previous statements on the nature of chain shifting (e.g., Łubowicz (2011), who states that all chain shifts, whether synchronic or diachronic, typically involve these properties). A property that is shared by these two kinds of shifts is the domain in which they occur, which is the segment. In this section, I argue that if we are to take the concept of chain shifting seriously, the first step in attempting to find a definition must be to restrict the domain of shifting to the segment. This was the tentative conclusion of Neasom (2011), from which I take updated versions of certain examples in what follows.

#### 5.3.1 Two kinds of counterfeeding, ‘environments’ and ‘sources’

Previous approaches to the domain of chain shifting can be split into roughly two camps, both of which pertain to the domain over which chain shifting can occur. That is to say, do both the  $A \rightarrow B$  and  $B \rightarrow C$  parts of the chain shift have to act on the same segment, or can they act on different segments? The first position, whose two most obvious standard-bearers are Moreton (2004a) and Łubowicz (2003a; 2003b, *et seq*), gives no explicit upper limit for the domain of chain shifting. Indeed, both authors make it clear that they consider the set of chain shifts and the set of counterfeeding processes to be essentially isomorphic. Consider the following two quotes, the first from the front matter of Moreton’s corpus, the second from Łubowicz (2003b):

“This is a collection of known or alleged instances of the phonological phenomenon variously known as:

- “counterfeeding”
- “synchronic chain shift”
- “underapplication”” (Moreton 2004a)

“[C]ounter-feeding opacity is a *chain-shift effect*” (Łubowicz 2003b, p.316).

It should be pointed out that Łubowicz splits her typology of chain shifts into ‘push shifts’, in which contrast is the only motivation for the  $B \rightarrow C$  part of the shift, and

‘regular shifts’ where there is independent motivation in the grammar for both the  $A \rightarrow B$  and  $B \rightarrow C$  shifts. Furthermore, her case study of a push shift (Finnish, from her 2003 thesis) has the domain of the segment, and the example she gives of a regular shift (Sea Dayak, in her (2011) review article) occurs across a larger domain. However, Łubowicz does not draw an explicit connection between size of domain and the classification of a particular shift as a push or a regular shift. In Neason (2013), I argued that the  $nt\# \rightarrow n\# \rightarrow \#$  shift in Catalan (discussed in the corpus in chapter 4) could be seen either as a regular or a push shift in Łubowicz’s preservation of contrast model, depending on one’s interpretation of the data. In sum, then, whilst Łubowicz makes a typological split within the class of chain shifts, the size of the domain does not appear to be the line along which the typological split is made.

On the other hand, there are approaches that explicitly state that the class of chain shifts is isomorphic not with the set of counterfeeding processes as a whole, but with a subset of counterfeeding processes. A common division made in the class of counterfeeding processes is between counterfeeding-*on-focus* and counterfeeding-*on-environment*. Counterfeeding-*on-focus* is defined by Baković, who is explicit that counterfeeding-*on-focus* and chain shifts are equivalent, in this way: “[I]f the two rules involved were ordered in such a way that Q precedes and thus feeds P then both rules would apply to the same segment” (2007, p. 222). He contrasts this with his definition of counterfeeding-*on-environment*: “if Q precedes and feeds P then both rules would still apply but this time not to the same segment” (ibid). In this case, then, it is obvious that what separates chain shifts from other counterfeeding processes is the domain of the effect. John McCarthy shares this view, defining a chain shift as a process in which “input *segment* /A/ becomes output [B] and input /B/ becomes output [C] in identical or overlapping environments” (2007a, p.129, emphasis mine).

It is clear that that this would substantially reduce the set of putative chain shifts, and would concomitantly reduce the power of any holistic theory of chain shifting. Indeed, I intend to argue in favour of the position that, if we are to properly investigate whether synchronic chain shift is a genuine phonological phenomenon, it will first be necessary to define synchronic chain shift as a sub-type of

counterfeeding-on-focus. Before moving on to discuss individual case studies I will briefly address some more general, conceptual issues, relating to two other metrics of relatedness within shifts.

Closely related to the domain of chain shifts is the environment in which both parts occur. Accounts based on the idea that chain shifts are instances of counterfeeding-on-focus also have tight restrictions on the environment of the processes that form the shift. McCarthy has long suggested that “traditionally” chain shifts have identical environments (1999, p. 365). However, it is also at least logically possible for a wider definition of chain shifting to require identical environments for the  $A \rightarrow B$  and  $B \rightarrow C$  parts of the shift. Consider the case in Catalan, for example, for which Moreton gives the schematic  $/nt\#/ \rightarrow /n\#/ \rightarrow /#\#/$ . It is possible to suggest that both the  $/t/ \rightarrow \emptyset$  and  $/n/ \rightarrow \emptyset$  parts of the shift simply occur word-finally, meaning that the environment for both is identical. This suggests that in order to make an argument that a domain restriction on chain shifts is necessary, it is also necessary to rule out the intermediate position that there can be genuine chain shifts above the level of the segment as long as the environment is identical. Later in this section I argue that there is no evidence that this intermediate theoretical position defines a useful class of processes.

A final important conceptual issue is whether there is a link between shifts at the level of the segment and shifts that have a single source (i.e., the motivation for the  $A \rightarrow B$  and  $B \rightarrow C$  parts of the shift is the same). There are certain chain shift theorists who remain agnostic on the size of domain that can be considered when discussing putative shifts, but feel more strongly about the necessity of a shared motivation for shifting. For example, David Mortensen draws a distinction between counterfeeding and chain shifting based on the intuition that “scalar” chain shifting processes “appear to arise from a single source” (2006, p.69). This divides the set of chain shifts in a completely different way to the approaches seen before. It suggests nothing about whether the size of domain involved is relevant, particularly as an effect like the Catalan shift definitely has a domain of more than one segment, but can be analyzed as being scalar (see Padgett (2002) and Neasom (2013) for discussion of this point).

The Catalan example also illustrates the difficulties inherent in defining what a “single source” is. Both steps of the shift,  $nt\# \rightarrow n\#$  and  $n\# \rightarrow \#$ , are instances of consonant deletion, suggesting a unified motivation. More general solutions like Moreton & Smolensky (2002) suggest that the  $A \rightarrow C$  mapping can be ruled out with a simple conjunction of MAX-seg, suggesting that the deletion processes are not different in a meaningful sense. However more detailed accounts of the effects (e.g., Mascaró (1976), Wheeler (2005)) show that the motivations, historical paths, and effects of the two instances of deletion are different. The literature on this point is voluminous, but a brief illustration of the differences in the two processes can be given in how productive each process is. Wheeler describes the cluster reduction in the  $A \rightarrow B$  part of the shift ( $nt\# \rightarrow n\#$ ) as “virtually categorical in Catalonia” (Wheeler, 2005, p. 222). By contrast, the  $n$ -deletion rule is riddled with exceptions. In fact, whilst it is reasonably common in the nominal paradigm, only two  $n$ -final verbs (*tenir* and *venir*) undergo the process. The amount of exceptions to  $n$ -deletion allow Wheeler to argue (in my view convincingly) that “there is no general *phonological* process of  $n$ -deletion” in Catalan (p.330).

On the other hand, at the level of the segment, the conceptual link between the steps of the shift often seems to be more obvious. A clear example comes from the set of metaphonic vowel raising effects, for example the Lena Spanish shift. To recapitulate once again, in the Lena Spanish shift, /a/ raises to [e], whilst /e/ raises to [i] before certain suffixes, all of which contain the high vowel {-i}. The conceptual similarity in this case is obvious; the same set of segments (vowels) undergoes the same kind of process (vowel raising) in the same context (/\_\_ {i}). Whilst there can be an argument over whether the two parts of the shift have differing histories, in the synchronic grammar it seems clear that the two processes are conceptually linked. In this section, I claim that there are no cases of chain shifting above the level of the segment that have a genuinely shared source, aligning myself with theorists like McCarthy and Baković, who view chain shifts as examples of counterfeeding-on-focus.

### 5.3.2 A particularly problematic case: Sea Dayak

I start with a case study that illustrates several of the problems inherent in attempting to classify interactions of processes above the level of the segment as chain shifts.

The effect, an interaction of cluster reduction and nasal harmony in Sea Dayak, was first reported in a brief discussion of the language by Scott (1957), who updated his remarks slightly in a later paper (1964) that compared the Sea and Land varieties of the Dayak language. The data given in Scott (1957, p.511) is shown below:

- (9) (a) nãŋãʔ ‘straighten’ vs. nãŋaʔ ~ nãŋgaʔ ‘set up a ladder’  
 (b) ramõʔ ‘timber’ vs. ramoʔ ~ ramboʔ ‘a kind of flowering plant’

The standard analysis (first made in Kenstowicz & Kisseberth (1979), though for a completely different take see Mielke et al (2003)) runs like this: Nasal harmony spreads from nasal consonants onto underlyingly oral vowels, but it cannot spread through an oral consonant. Independently of this, underlying stops in homorganic nasal-stop sequences are optionally deleted. This has been claimed to be a chain shift because of cases like *ramoʔ*. In this case, it appears that nasal harmony has underapplied, because there is a nasal consonant adjacent to an oral vowel. The reason for this is that nasal harmony applies before cluster deletion. Therefore, even though the nasal is in a position to spread its nasality, it is too late in the derivation. Kenstowicz and Kisseberth model the problem with an extrinsic, counterfeeding order, as shown in table 5-4.

Table 5-4: The Kenstowicz & Kisseberth analysis of Sea Dayak

	/naŋaʔ/	/naŋgaʔ/
Nasal Harmony	nãŋãʔ	nãŋgaʔ
Cluster Reduction	-	nãŋa
	[nãŋãʔ]	[nãŋaʔ]

Importantly for the rest of this section, the Sea Dayak effect is listed in Moreton’s corpus (2004a). The schematic given for the effect is ŋga → ŋa → ñã. In a later subsection, I will take issue with this schematic, arguing that it is not genuinely coherent, but for now I move on to discuss more general issues with the Sea Dayak data, and previous approaches to it.

### 5.3.2.1 Problem 1: The data itself

This data is, putting it mildly, somewhat scant. However, just the two minimal pairs have been used as test cases or data points in favour of at least four theoretical approaches of counterfeeding interactions (Kenstowicz & Kisseberth (1979), Moreton & Smolensky (2002), Mielke et al (2003), Ettliger (2007), and Łubowicz (2011)). In all of these sources the optionality of the cluster reduction rule is either ignored totally, or marginalized. It is also worth noting that the descriptions of the effect given by Scott suggest certain additional properties, chiefly gradience:

“The plosives and affricates and the sibilant may be preceded by a homorganic nasal... The voiced plosives and affricates are often very gentle in this case, and the distinction by ear of *rambuq* ‘a kind of flowering plant’, and *ramuq* ‘timber’, may depend mainly on the absence of nasality from the final vowel of the first word, thus **ram(b)oʔ, ramōʔ**” (1957, p.511).

“In Sea Dayak, the plosive is often very weak or even absent, but the absence of nasality in the following vowel indicates that what precedes is not regarded to be a simple nasal” (1964, p.433).

It is clear from the quotations above that there are two salient properties of the homorganic cluster deletion rule. The first is that it is optional. In some cases, there is a plosive. Whether or not it is weak is not important in deciding whether the rule is optional or not. If there are any cases in which the plosive is present, the rule cannot be said to be obligatory. The second property of the rule is that it is gradient. Both of Scott’s quotes suggest that it is not a simple matter of the obstruent being realized or not. In many cases, a version of the obstruent is realized, but one that is weaker than one would ordinarily expect. At this point, I state only that a rule that is based on just two minimal pairs, and that seems to operate both optionally and in a gradient manner is perhaps not the most robust example of counterfeeding. The optionality and gradience will take on greater significance in a subsequent subsection, where I argue that, even if one wishes to postulate an  $A \rightarrow B \rightarrow C$  order for the Sea Dayak effect, it falls out of the general principles of any version of the phonological grammar that includes a lexical and a postlexical component (contra Ettliger 2007).

5.3.2.2 Problem 2: Counterfeeding/Chain Shift analyses of the data are not insightful

The data itself, then, appears to be a somewhat shaky foundation on which to test a theoretical principle. However, we must assume that it is at least possible that the deletion rule is reasonably regular and reasonably productive across the language, even if it is optional and gradient in each individual instantiation of the rule. It is therefore important to look into any theoretical approach that treats the lack of an A → B → C mapping as significant. The two main approaches that do this are Kenstowicz and Kisseberth's counterfeeding account, and a Local Conjunction account offered in Łubowicz (2011). It is important to state that Łubowicz does not have a theoretical commitment to Local Conjunction (though she is not against it per se; see Łubowicz 2005), and that she is merely showing how an LC account of Sea Dayak *would* work, were one minded to do so. However, she is explicit that she does consider the effect in Sea Dayak to be a chain shift. The LC analysis, following Łubowicz (2011) (the tableaux are adapted from p.1723), is shown in table 5-5.

Table 5-5: Slightly adapted tableaux for Sea Dayak

/naŋa/	[IDENT[nas] & MAX] <sub>AdjSeg</sub>	*ND	*NV	IDENT[nas]	MAX
nãŋga		*!		(*)	
nãŋa			*!	(*)	
☞ nãŋã				(*)*	

/naŋga/	[IDENT[nas] & MAX] <sub>AdjSeg</sub>	*ND	*NV	IDENT[nas]	MAX
nãŋga		*!		(*)	
☞ nãŋa			*!	(*)	*
nãŋã	*!			(*)*	*

The constraint that forces vowel harmony is \*NV, in which a violation is assigned for every output sequence of a nasal consonant followed by an oral vowel. Cluster reduction is forced by \*ND, which assigns a violation to any sequence of a nasal consonant followed by an obstruent. These markedness constraints outrank the two

faithfulness constraints IDENT[nas] and MAX, the first of which works to preserve the underlying nasality value of each segment, whilst the second disallows deletion. These constraints, in any Markedness >> Faithfulness order, are enough to get the correct output form, [nãŋã] from the input /naŋa/, but if the underlying input is /naŋga/, then this ranking will give [nãŋã] as the optimal output, because it satisfies all of the markedness constraints. The solution, in a Local Conjunction analysis, is to conjoin the two faithfulness constraints IDENT[nas] and MAX. As the unwanted candidate *nãŋã* violates both of these constraints, and neither of the other input candidates above do, only this candidate triggers the highly-ranked conjoined constraint.

This account is problematic, because the rationale for postulating this kind of conjunction is essentially circular. The reason that we conjoin MAX and IDENT[nas] appears to be that only candidates exhibiting violations of both of these features are ruled out in the data. When we ask why it is that they are ruled out in the data, all that we are able to say, from the Local Conjunction perspective, is that they violate the super-constraint [IDENT[nas] & MAX]<sub>AdjSeg</sub>. It can be argued that, like all conjoined constraints, the idea is to keep inputs and outputs as minimally different from one another as possible, but without additional reasons to use Local Conjunction, and not simply the existing markedness and faithfulness constraints already in play in the language, this argument does not hold much force, as it seemingly draws a connection between two properties (nasality and the presence vs. absence of a segment) that have no clear functional relation. The question that this immediately raises is whether the two processes involved in the Sea Dayak chain shift are in any way genuinely related, which is discussed in more detail below.

### 5.3.2.3 Problem 3: The schematic

Consider the two schematics below, both from Moreton (2004a). (10a) is the schematic for Lena Spanish (slightly adapted, as in Moreton's corpus it is incorrectly listed as being a lowering shift of the form  $i \rightarrow e \rightarrow a$ ). (10b) is the schematic for Sea Dayak:

- (10) (a) Lena Spanish: stressed  $a \rightarrow e \rightarrow i$  before u  
 (b) Sea Dayak:  $\eta ga \rightarrow \eta a \rightarrow \eta \tilde{a}$

There are two key differences here, both of which suggest that the A → B and B → C steps are intrinsically related in the Lena shift, but not in the Sea Dayak case. The first is the amount of segments in each step in the shift. In the Lena case, there is only one segment involved at each stage of the shift, whereas there are multiple segments involved at all stages in the Sea Dayak schematic. Indeed, the A part of the shift involves three segments, whilst the B and C parts of the shift involve two. The reason for this lies in the second difference between the two schematics. In the Lena case, Moreton places information about the environment in which vowel raising takes place outside of the schematic itself, with the simple declarative statement ‘before u’. Therefore, the schematic includes *all and only what changes* as a result of the chain shift, in this case the vowel raising process. This is not the case in the Sea Dayak schematic, which shows not only what changes, but also builds environmental information directly into the schematic itself. This, in my opinion, creates the illusion of coherence where none genuinely exists.

To illustrate this, we can show through simple linear rules what is actually happening at each stage of the shift. The A → B part of the shift is the cluster deletion rule:

$$(11) [-\text{son}, -\text{cont}] \rightarrow \emptyset / [+nas, +\text{cons}] \_ \_$$

After a nasal consonant, a stop is optionally deleted. The B → C part of the shift is the vowel harmony rule, which can be written like this:

$$(12) V \rightarrow [+nas] / [+nas] \_ \_$$

It should be pointed out that the environment for the two processes is not completely shared. Indeed, the environment for the cluster deletion rule is a subset of the environment for the rule of vowel harmony. This can be shown by further examples from Scott (1957, p.511):

- (13) (a) m̄ata ‘an eye’  
 (b) m̄õã ‘the face’  
 (c) m̄ajã ‘a season’

From (13a), we can see that the cluster deletion rule must be explicit that only nasal consonants condition the deletion. If plosives appear as singletons, and are preceded by a nasal *vowel*, deletion does not occur. This suggests that it is actually not the nasality of the first consonant in a C<sub>1</sub>C<sub>2</sub> cluster that is the crucial factor in the deletion, but rather the CC nature of the cluster itself. Examples (13b) and (13c) show that spreading from one vowel to another is possible (the glide in (13c) is transparent for nasality, as spread is only blocked by [+cons] segments), so the only feature required in the environment is [+nas]. Any further specification predicts that nasal harmony will spread less far than it actually does. If we ignore this - though I would argue that we should not, as it offers further evidence that the two processes are unrelated - then we can begin to make the Sea Dayak schematic in (10b) analogous to the Lena schematic in (10a), by listing *only and all of the segments that change*, along with a simple statement of the environment in which the changes occur. The original schematic is shown here for clarity:

- (14) ŋga → ŋa → ŋã

If we take the environmental information out of this schematic, we are left with the altered version below:

- (15) ga → a → ã / [+nas, +cons] \_\_

I would argue that this still presents a somewhat misleading picture of the Sea Dayak effect. This is because, in the A part of the shift, the /a/ that follows the /g/ is not necessary. Sea Dayak, on Scott’s analysis, does not have final ND clusters, so we can assume that the cluster deletion operates at least vacuously, even when a vowel does not follow. This means that the true schematic does not have a coherent A → B → C structure:

- (16) [-son, -cont] → Ø / [+nas, +cons] \_\_, V → ~V / [+nas] \_\_

In (16), the independence of the two processes is fully laid bare. There is nothing shared in the structural description of the two rules. One acts on consonants and the other acts on vowels. These two rules only interact at all because there is one context in which both operate: after a nasal consonant. However, there are contexts where the nasal harmony rule operates, but the cluster deletion rule does not.

In sum, then, whilst it appears clear that both a straightforward counterfeeding order and a local conjunction analysis can both describe the interaction we observe in Sea Dayak, neither is really able to explain it. It is surely worth asking whether there is a good reason to suppose that cluster reduction should follow nasal harmony, if we are to think about the problem in terms of a serial derivation, or whether there is a genuine functional link between nasality and the presence or absence of segments, if we are to take a parallel approach. In the next subsection, I argue that (contra Ettliger (2007)) the data suggests that cluster reduction is certainly a phonetic, and perhaps a postlexical process. If we assume that nasal harmony applies within the lexicon<sup>24</sup>, then the  $A \rightarrow B \rightarrow C$  ordering falls out as a necessary consequence of any theory of grammar that includes lexical and postlexical components.

#### *5.3.2.4 Two intrinsically ordered accounts of Sea Dayak*

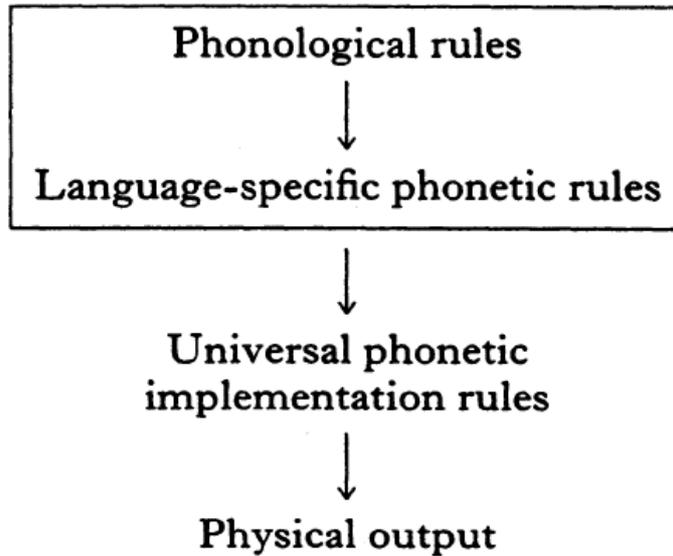
As I mentioned earlier in this chapter, Kenstowicz & Kisseberth (1979) model the Sea Dayak effect through a straightforward counterfeeding order, where the  $B \rightarrow C$  rule precedes the  $A \rightarrow B$  rule. My assumption is that for proponents of a serial, rule-based grammar (see Vaux (2008) for a robust defence of this position), there is no real reason to change this. The problem with this account for many is that the ordering appears to be extrinsic: is there a good reason to suggest that the  $A \rightarrow B$  rule should come later in the derivation? I would argue that, at a basic level, there is, if we take the position that it is a rule of phonetic implementation rather than a genuinely phonological rule. Late phonetic rules have been a part of rule based phonology since at least *SPE*, and are retained in more modern rule-based approaches, like Lexical Phonology (Kiparsky 1985)/. As shown in figure 5-3, first

---

<sup>24</sup> We have no knockdown evidence for this from Scott's work, but equally there is no suggestion that it applies across words, and reports from related languages suggest that nasal harmony is lexical (see, for example, Cohn (1993, p.54ff) on Sundanese).

published in Cohn (Cohn, 1993, p. 44), models of phonology which allow for computation on multiple levels order genuinely phonological rules before phonetic rules, whether those rules are language-specific or universal:

*Figure 5-3: The organization of the phonological grammar*



We do not know much about the process of cluster reduction in Sea Dayak. The one thing that we do know with some degree of confidence, as it is mentioned in both of Scott's discussions of the effect, is that the process is essentially gradient in character. Sometimes full deletion occurs, sometimes the stop is produced, but more softly than would ordinarily be expected. A major distinction that Cohn (following Keating (1988; 1990) and Pierrehumbert & Beckman (1988)) draws between phonological and phonetic processes is that phonological processes are generally categorical, whilst phonetic processes are usually gradient. Scott does not give any information about whether nasal harmony is categorical or gradient in Sea Dayak, but in the absence of discussion, or detailed instrumental data, the default assumption must be that the process is categorical. If this is true, then the ordering relations in the Sea Dayak case turn out to be both simple and principled.

Table 5-6: Phonology: Nasal Harmony applies wherever possible

	/naŋa/	/naŋga/
Nasal Harmony	nãŋã	nãŋga
Intermediate representation	nãŋã	nãŋga

Phonetics: Cluster reduction applies (in the table below, it applies fully, but this is not necessarily the case):

	nãŋã	nãŋga
Cluster Reduction	n.a	nãŋga
Surface representation	[nãŋã]	[nãŋga]

I make no particular commitment to how each of these processes should best be represented. The salient point is simply that, if the grammar orders phonological rules before phonetic rules, the  $A \rightarrow B \rightarrow C$  order of the effect in Sea Dayak is not only unremarkable, but unavoidable.

If one does not wish to draw a distinction between phonological and phonetic rules, vis-à-vis where they can be ordered, then of course this analysis is of no help. However, as Cohn notes, there are striking similarities in the distinction between phonological and phonetic rules on the one hand, and lexical and postlexical rules on the other. Cohn’s position is that “[a]lthough the distinction is mainly terminological, I take the position that postlexical phonological rules and phonetic implementation rules are distinct” (1993, p.46). The account of the Sea Dayak effect that I have given above is in line with this distinction, but there is also an account that would be in line with cluster reduction as a postlexical phonological rule.

In what follows, I use a simplified version of Stratal Optimality Theory (see Bermudez-Otero (2011) or Kiparsky (2015) for a concise introduction) as my method of representing the principles of Lexical Phonology. This is because it is more current than rule based Lexical Phonology approaches, but is based on the same guiding principles (chiefly from Kiparsky (1985; 2000)). Stratal OT makes the assumption that there are at least two separate strata of phonological computation, one occurring at the lexical level and one at the post-lexical level. Kiparsky posits that standard accounts of Optimality Theory cannot model opaque effects largely because of their parallel nature. For Kiparsky, it is impossible for the modelling of

opaque processes to proceed without some kind of serial derivation. He suggests that this can be best-achieved, within the OT framework, via a framework which includes separate levels of computation for lexical and postlexical processes. On this kind of account, opaque effects fall out naturally from the organizational principles of the grammar, as a result of what Kiparsky calls “inter-level constraint masking” (Kiparsky 2000, p.351). That is to say, the effect of a constraint at one level may be obscured or undone by what occurs on a following level.

Ettlinger (2007) discusses a Stratal OT account of Sea Dayak, stating that it cannot work because “there is no evidence that nasal harmony and post-nasal stop deletion are anything other than word-level processes because they both apply to the whole word” (p.7 of online version)<sup>25</sup>. It is true that Scott’s descriptions are of individual words, so it is not possible to see whether either rule applies at the phrasal level. What little direct evidence one can adduce from the primary literature, however, suggests that cluster reduction may be a post-lexical process. Not only is it gradient, as discussed in the previous section, but it appears to be an optional process as well. As there is no evidence that nasal harmony is anything other than categorical and obligatory, I believe that there is at least suggestive evidence that nasal harmony can be placed within the lexical phonology (at the word-level, as Ettlinger suggests) and cluster reduction in the post-lexical phonology. If this step is taken, then a Stratal OT account becomes straightforward.

(17) Constraints (adapted from Łubowicz (2011) and Ettlinger (2007)):

- (a) \*NV<sub>Oral</sub>: Assign a violation for any oral vowel directly following a nasal
- (b) \*ND: Assign a violation for any voiced obstruent directly following a nasal consonant
- (c) IDENT[nas]: Assign a violation for any output segment that differs from its input correspondent in its specification for nasality
- (d) MAX: Assign a violation for any input segment that does not have an output correspondent (no deletion)

---

<sup>25</sup> It is worth noting in passing that Ettlinger’s account of Sea Dayak, which is based on exemplar theory, is also a non chain-shift analysis. I do not discuss it at length for reasons of space, and because it necessitates a discussion of exemplar theory which falls beyond the scope of this thesis.

Table 5-7: Stratum 1: Desired intermediate outputs: /naŋa/ → |nãŋã|, /naŋga/ → |nãŋga|

Constraint ranking: MAX >> \*NV<sub>ORAL</sub> >> \*ND >> IDENT[nas]

/naŋa/	MAX	*NV <sub>ORAL</sub>	*ND	IDENT[nas]
nãŋa		*!		*
nãŋga			*!	*
☞ nãŋã				**

/naŋga/	MAX	*NV <sub>ORAL</sub>	*ND	IDENT[nas]
nãŋa	*!	*		
nãŋã	*!			
naŋga		*!	*	
nãŋgã			*	**!
☞ nãŋga			*	*

In this first stratum, the lexical process of nasal harmony applies. Because there is no deletion at this stage in the derivation, MAX is highly ranked. The constraint enforcing nasal harmony, \*NV<sub>ORAL</sub>, is also highly ranked. In neither of the optimal outputs is there any deletion, nor is there any instance of a nasal consonant adjacent to an oral vowel. Note that \*ND must outrank IDENT[nas], otherwise the optimal output for the input candidate /naŋa/ would be [nãŋga], with an epenthetic [g] inserted.

In the second stratum, the constraint mandating that nasal consonants spread their nasality onto following vowels is demoted, along with MAX, and \*ND, a constraint that rules out nasal-consonant clusters, is promoted.

Table 5-8: Stratum 2: Desired outputs: |naŋa| → [nãŋã] → [nãŋã], |naŋga| → [nãŋga] → [nãŋa]

nãŋã	*ND	ID[nasal]	MAX	*NV <sub>ORAL</sub>
nãŋa		*!		*
nãŋga	*!	*	*	
☞ nãŋã				

nãŋga	*ND	ID[nasal]	MAX	*NV <sub>ORAL</sub>
☞ nãŋa			*	*
nãŋã		*!	*	
naŋga	*!	*		*
nãŋgã	*!	*		
nãŋga	*!			

The two highly ranked constraints are not ranked with respect to one another, and neither are the two low-ranked constraints. The only crucial ranking is between the two groups. It is vital that \*ND, ID[nasal] >> MAX, \*NV<sub>ORAL</sub>, as for both inputs the optimal candidate violates one or both of these lower ranked constraints.

The overall point of this analysis is to show that, if simple assumptions are made about Sea Dayak phonology, we can rely on machinery that is already present in many theorists' version of the grammar to give the correct, allegedly chain shifting order purely through the basic architecture of either the phonological grammar as a whole, or Lexical Phonology. In both cases, the ordering of the two processes is not extrinsic, and there is no suggestion that they are related processes, or that this ordering in any way reflects a relationship between the two specific processes. This is very different to the Local Conjunction analysis adopted in Łubowicz (2011). If we are to take constraint conjunction seriously, then there must be some principled reason to conjoin the particular constraints that make up the super-constraint. Otherwise, local conjunction amounts to little more than data-fitting.

The wider issue that this raises is about the nature of chain shifting. Chain shift theories are built on the notion that the lack of an  $A \rightarrow C$  mapping is problematic, and some bespoke representational or derivational device is necessary to stop this from happening. In this sense, there is something important about the  $A \rightarrow B \rightarrow C$  mapping that is a feature of chain shift. This section has argued that, in Sea Dayak at least, the  $A \rightarrow B \rightarrow C$  mapping is illusory. Even if it is argued that there is still some semblance of an  $A \rightarrow B \rightarrow C$  mapping, its existence is completely epiphenomenal.

The seeming interaction of nasal harmony and cluster reduction happens in the way it does because of substantive differences between the processes themselves.

In the following sections, I simultaneously pursue two distinct arguments. The first is the section internal argument that seeming chain shift processes that exist above the level of the segment should not be viewed as chain shifts. The next subsection is a further example of the problematic cases that we have to admit if we allow chain shifting above the level of the segment. The second, more general point is that seeming chain shifts are often epiphenomena of other, better attested processes. It is this idea that I discuss in the final sub-section of this part of the chapter.

### 5.3.3 *No shared environment, no $A \rightarrow B \rightarrow C$ mapping: Icelandic*

As another demonstration of the lack of coherence that is exhibited in discussions of chain shifts above the level of the segment, it is worth considering an example from Icelandic that appears in Moreton's corpus, taking the schematic form  $aCr\# \rightarrow aCur\# \rightarrow \ddot{o}Cur\#$ . This effect is an interaction between a rule of *u*-umlaut that is discussed in various analyses of Icelandic (see, e.g., Orešnik (1972) Kiparsky (1985), Árnason (1988; 2011), Thráinsson (1994), Kenstowicz (1994), Gibson & Ringen (2000)), and a separate rule of epenthesis. The data below, taken from Kenstowicz (1994, p.80), illustrates the pattern:

(18)	ACC. SG	NOM. SG	DAT. PL	GLOSS
	(a) hatt	hatt-ur	hött-um	'hat'
	(b) dal	dal-ur	döl-um	'valley'
	(c) stað	stað-ur	stöð-um	'place'

The majority of morphological paradigms in Icelandic work like the dative plural. Before suffixes featuring /u/, underlying /a/ fronts to surface [ö]. However, in the nominative singular, this fails to happen. The simplest reason for this, and the one explored in Kenstowicz, is that the [u] vowels in the nominative singular forms are epenthetic. This epenthesis can be ordered after the umlaut rule, thus explaining why epenthesis fails to apply in these instances. It should be pointed out that this is not the only way of dealing with this problem. Nor is it entirely clear that there is a problem here to be dealt with. Sources differ on how active the *u*-umlaut rule is in

Icelandic<sup>26</sup>. Árnason (1988), for example, seems uncomfortable with the idea of *u*-umlaut as an active process:

“Generally, the formations where umlaut does not occur are recent ones, and this strongly suggests that *u*-umlaut is not ‘active’ in word formation” (p.7)

“the question of whether *u*-umlaut is in some sense still ‘active’...cannot be answered” (p.13)

On the other hand, Thráinsson (1994) is very clear that he does believe that *u*-umlaut is an active part of the phonological grammar, to the point of directly contradicting Árnason:

“That this rule is alive and well can be seen from the fact that it applies in new words, (inflected) loan words and even foreign names that are inflected” (p.152)

Given that I am certainly in no position to adjudicate on whether *u*-umlaut is active or not in the phonology of Icelandic, I let the matter drop at this point, and take my general attitude that we should, for now at least, assume that the process is active. As long as it remains a possibility that there is some synchronically active process of *u*-umlaut in Icelandic, it requires some explanation. I would argue, however, that whatever form this explanation takes, the Icelandic effect should not be considered a chain shift. This is because it has neither a coherent  $A \rightarrow B \rightarrow C$  mapping, nor a shared environment for the putative  $A \rightarrow B$  and  $B \rightarrow C$  processes to take place in.

To make sense of both of these claims, it is useful to first show what the processes of epenthesis and *u*-umlaut look like when cast in a simple, linear-rule format:

- (19) (a) Epenthesis ( $A \rightarrow B$ ):  $\emptyset \rightarrow u / C\_r$   
(b) *U*-umlaut ( $B \rightarrow C$ ):  $a \rightarrow \ddot{o} / \_C_{(0)}u$

Turning first to the environment, we first see that there is no unified statement that can be made, except for the somewhat trivial claim that a consonant may be involved

---

<sup>26</sup> Thanks are due to Renate Raffelsiefen for suggesting to me that *u*-umlaut may not be a phonologically active process, and for making me read more about it!

in both parts. The epenthesis rule takes place in order to break up a consonant cluster, whilst the only crucial piece of environmental information for the *u*-umlaut rule is contained in the name of the process – before /u/, umlaut occurs. This is, to my mind, the crucial difference between a shared context and a shared environment. When the parameters of what can be shared are as loose as they are in the Icelandic case, then any number of processes may well come to interact.

There is nothing inherently interesting about this, no reason to search for some higher explanation for why this kind of interaction should occur. In cases where there is a shared environment, where both processes happen in exactly the same place and nowhere else, there is a motivation either to collapse the two rules into one, unified process, or at least relate the two processes by giving them the same motivation. The first approach is common in accounts of, for example, metaphony, where both  $A \rightarrow B$  and  $B \rightarrow C$  mappings can be seen as instances of the same overall process. The second is typical of lenition analyses, where the two processes are different in terms of their effects, but the same in terms of their overall motivation. In the interaction between epenthesis and *u*-umlaut, however, it is clear that there are two completely separate motivations at work. Epenthesis is employed to break up an unwanted cluster of consonants, which is a totally implausible motivation for umlaut, a process motivated in this instance by the presence of certain suffixes containing /u/.

Further to this, if we examine again the segments that are affected by each of the rules, and assume that, as in the Sea Dayak case study, we should strip out any unnecessary information from the schematic, we are left with a set of mappings that do not show anything that is recognizable as a chain shift. Consider the schematic from Moreton's corpus:

$$(20) \quad aCr\# \rightarrow aCur\# \rightarrow \ddot{o}Cur$$

For the operation of the first rule, the /a/ is completely unimportant. This epenthesis would occur whatever the nature of the preceding vowel. This means that, in terms of the schematic, the /a/ serves no purpose and can be discarded. As well as this, the  $Cr\#$  part of the schematic is information about the environment of the effect, not the segments which undergo any change. If we are to have a consistent system of

schematizing chain shifts, where *all and only* what changes is included, this must be discarded as well, meaning that the first part of the schematic should be:

$$(21) \quad \emptyset \rightarrow u$$

The B part of the schematic is just as problematic. In this instance, the only thing that changes is underlying /a/. This means that *everything* that follows it in Moreton's schematic (Cur#) is either unnecessary (C, r) or environmental information (u, #). Again, following the dictum laid out previously, this would suggest the following schematic for the Icelandic 'shift':

$$(22) \quad \emptyset \rightarrow u, a \rightarrow \ddot{o}$$

As mentioned above, the problem in Icelandic runs even deeper than the problem in the previous subsection, because there is no way of postulating a coherent environment for these two processes, unlike in Sea Dayak. This, then, is a completely random coming together of two processes. It may even be the case that there is no coming together at all. The umlaut rule appears to be morphologically conditioned, rather than always needing an overt phonological trigger. The evidence for this is instances of the application of umlaut without a triggering *u*. The examples in (23) illustrate this (all from Árnason (2011, p. 244)):

(23)	SING.	PL.	GLOSS
(a)	b[a]rn	b[ö]rn	'child'/'children'
(b)	fj[a]ll	fj[ö]ll	'mountain'/'mountains'
(c)	bj[a]rg	bj[ö]rg	'cliff'/'cliffs'

It may well be that there is nothing about the phonological properties of the /u/ that causes the umlaut, but rather the set of affixes (or a zero-affix, in the case of the plural). If this is so, then the failure of forms featuring epenthetic [u] to undergo the umlaut rule makes perfect sense; the environment for umlaut to occur simply never applies. The nominative singular is not a morphological environment for umlaut. Thus, whether or not there is an [u] in such forms is completely irrelevant. The rules do not even touch.

In each of the cases above, even if one does not buy into the specific arguments about potential non chain-shift solutions for these effects, what should be clear is that there are two processes at work. These are examples of what Łubowicz would call ‘regular shifts’. Perhaps then, the argument is not about the size of the domain, but about unity of purpose. The two examples above do not allow us to tease this distinction apart, as it is clear that there is no such unity in either case. It is important, then, to discuss a shift that takes place at above the level of the segment, but in which both parts of the shift occur for the same reason. The sub-section below addresses this question with an analysis of the Siouan language Hidatsa.

#### *5.3.4 When unification gets too unified: A seeming vowel deletion chain shift in Hidatsa.*

In the previous section I argued that, in cases like the Sea Dayak and Icelandic effects, the seeming  $A \rightarrow B \rightarrow C$  mapping is an unexceptional, epiphenomenal consequence of the way that the grammar is arranged. An analysis of Sea Dayak makes *more* sense if we place the processes that comprise the  $A \rightarrow B$  and  $B \rightarrow C$  mappings on different levels, whether those levels be phonology vs. phonetics or lexical vs. postlexical. In this section, I discuss an effect in Hidatsa where this kind of analysis would make no sense at all; the process that leads to the  $A \rightarrow B$  and  $B \rightarrow C$  mappings is exactly the same in both cases, meaning that there can be no principled separation across levels. My argument is that the  $A \rightarrow B \rightarrow C$  mapping is just as unexceptional and epiphenomenal in this case, just in a different way. In this case, the process that applies is indifferent to whether the structure it is applying to is an example of underlying /A/ or underlying /B/.

I have discussed the conceptual problems with this idea, with reference to Hidatsa, as well as two other deletion shifts in Catalan and Chemehuevi, in previous work (Neasom (2011; 2013)). What follows is not entirely dissimilar from those papers, but is significantly upgraded.

In Hidatsa, there is a productive process of what appears to be subtractive morphology. In the earliest extant analysis of this effect, Zellig Harris proposes a rule whereby the final mora of a word is deleted to form the imperative (1942). Some

examples are given in (24), where I list the original sources. I should note that these sources come from Zimmermann's (2014) thesis, which also discusses Hidatsa.

(24)	STEM	IMPERATIVE	GLOSS	SOURCE
	<i>A → B</i>			
(a)	kikua	kiku	'set a trap'	Z. Harris (1942, p.171)
(b)	ika:	ika	'look'	Z. Harris (1942, p.171)
(c)	na:	na	'go'	Boyle (2007, p.202)
(d)	wia	wi	'cry'	Boyle (2007, p.202)
(e)	kuraʔa:	guraʔa	'carry'	Boyle (2007, p.201)
	<i>B → C</i>			
(f)	cixi	cix	'jump'	Z. Harris (1942, p.171)
(g)	ra:pa	na:b	'pass by'	Boyle (2007, p.202)
(h)	awa:ki	awa:g	'sit'	Boyle (2007, p.202)

The conception of this process as a chain shift in Moreton & Smolensky (2002) rests on the idea that long vowels become short as a result of this process, whereas short vowels are deleted entirely. This is the intuition behind Moreton's schematic, which is  $V_1V_2\# \rightarrow V_1\# \rightarrow \#$ . Moreton & Smolensky suggest that the process can be modelled via self-conjunction of MAX. The idea behind this is that the lack of an  $A \rightarrow C$  mapping (in this case the complete deletion of underlying long vowels) is problematic. This allows for a constraint that heavily penalizes extreme deletion, whilst remaining relatively sanguine about the deletion of a single mora. This analysis is shown below. It is, I should stress, my imagining of how Moreton & Smolensky's analysis would work, as they do not provide tableaux for the Hidatsa effect.

(25) Constraints (following Moreton & Smolensky (2002) (for the MAX's) and Wolf (2011) (for the markedness constraints)).

- (a) \*HIATUS = Assign a violation for any instance of hiatus
- (b) FINAL-C (following, e.g., McCarthy (1993)) = Assign a violation for any word-final vowel
- (c) MAX = Assign a violation for any deleted segment
- (d) MAX & MAX<sub>AdjSeg</sub> = Assign a violation for any two violations of MAX that occur in adjacent segments

Table 5-9: Local Conjunction tableaux for Hidatsa

/kikua/	MAX & MAX <sub>ADJSEG</sub>	*HIATUS	FINAL-C	MAX
kikua		*!	*	
☞ kiku			*	*
kik	*!			**

/cixi/	MAX & MAX <sub>ADJSEG</sub>	*HIATUS	FINAL-C	MAX
cixi			*!	
☞ cix				*
ci	*!		*	**

In the tableau for /kikua/, which represents the A → B part of the shift, the fully faithful candidate is ruled out by a markedness constraint. In this case, I have followed Wolf's (2011) template for any instance of a self-counterfeeding process. Wolf lists Hidatsa and Catalan as instances of such processes, and his analysis (which is couched in OT-CC, and refers to such processes as chain shifts) requires a constraint against hiatus, on the assumption that hiatus avoidance is a crosslinguistically plausible basis for a constraint that would penalize a sequence of two vowels. The candidate in which both vowels are deleted, [kik], is ruled out by virtue of its two violations of MAX, which in turn trigger a violation of the highly ranked super-constraint MAX&MAX<sub>AdjSeg</sub>. Therefore, the correct output candidate [kiku] is chosen.

In the tableau for /cixi/, which represents the B → C part of the shift, the fully faithful output candidate [cixi] is ruled out, in favour of the consonant-final form [cix], by a constraint against vowel final outputs, FINAL-C (see McCarthy (1993), Wolf (2011)). At this point, it is worth raising the issue that the constraints against hiatus and final consonants form two separate motivations for deletion. This is inherently problematic given what we know about the deletion process, i.e., it takes place word-finally, and it has a unified morphological function (forming the imperative).

This unified morphological function suggests that we need to think carefully about whether there is actually some pressure against word-final vowels in the language of Hidatsa as a whole. There is evidence from fieldwork on the language by Boyle (2002; 2007; No Year), that there is no general proscription against either vowel-final words or, perhaps more damningly, vowel hiatus in the language. The data in (26a-26k) illustrates words in which both of those structures surface.

- (26) *VV#*
- |                             |                                 |
|-----------------------------|---------------------------------|
| (a) <i>nii waara wa taa</i> | ‘are you OK?’ (online)          |
| (b) <i>cagii-ha hah kuu</i> | ‘live in a good way’ (online)   |
| (c) <i>axbishahbua</i>      | ‘seventeen’ (online)            |
| (d) <i>áaciwiricaraa</i>    | ‘milk lard/cream’ (2002, p.109) |
- V#*
- |                              |  |
|------------------------------|--|
| (e) <i>nii doosha</i>        | ‘how are you?’ (online)                              |
| (f) <i>rupá</i>              | ‘two’ (2002, p.103)                                  |
| (g) <i>arukirahpíreešawa</i> | ‘there was not a place to get down’<br>(2002, p.104) |
| (h) <i>wašúkawa</i>          | ‘the dog’ (2002, p.105)                              |
| (i) <i>akuhiši</i>           | ‘one which is red’ (2002, p.105)                     |
| (j) <i>aruhiši</i>           | ‘the red part (of something)’<br>(2002, p.105)       |
| (k) <i>istá</i>              | ‘his/her eyes’ (2002, p.109)                         |

If Moreton & Smolensky's analysis were applied to any of the words above, then the incorrect output form would be selected. If the constraint ranking is kept constant from that which is required to model the deletion effect, it is only to be expected that it will select incorrect outputs if final vowel-sequences or vowels are required. Fortunately, there is a simple explanation for why we sometimes observe deletion and sometimes retention of vowels, if we simply take on board Harris' original insight that the reason deletion occurs is for exponence of morphological information rather than any phonological pressure.

There is an immediate consequence of this realization for the notion of Hidatsa as a chain shift. If the effect is morphological, then there is no reason to suppose that it should ever apply twice. In simple, linear terms, the rule for deriving the imperative can be written out as shown below (adapted from Harris (1942)):

$$(27) \mu \rightarrow \emptyset / \_ \#$$

The only crucial element of the environment of the rule is that it occurs word-finally. It is *completely unimportant* what precedes the deleted material. The rule applies in exactly the same way whether what precedes the final vowel/mora is another vowel or a consonant. It is thus more insightful to schematize the effect in Hidatsa as  $V\# \rightarrow \emptyset$  rather than the  $V_1V_2\# \rightarrow V_1\# \rightarrow \#$  suggested in Moreton's corpus. This is directly analogous to my comments about the schematic in Sea Dayak and Icelandic in the previous section.

An argument that one could make against this position is that I am perhaps being something of a Luddite, insisting on a simple rule and ignoring more modern methods. However, there are state of the art approaches to similar problems that preserve the important insight that the motivation for the putative  $A \rightarrow B$  and  $B \rightarrow C$  parts of the shift is the same. Zimmermann (2014, pp.307-309) analyses imperative formation in Hidatsa as the attachment of a morpheme that creates an illicit structure. Because this structure is not an allowable surface form, and as a function of the way other constraints are ranked in Hidatsa, the net result of this affixation is, somewhat counter-intuitively, deletion. This is similar in spirit to Bye & Svenonius' (2012) account of a similar effect in Tohono O'odham, where the perfective is formed by

the deletion of a word-final consonant. Again, the process is one of affixation that creates a structure so objectionable to the phonology of the language that deletion is the final result.

These more abstract treatments have two crucial advantages over Moreton & Smolensky and Wolf's analyses:

- (28) (a) They do not affect forms that are not in the relevant morphological paradigm
- (b) They do not require any special mechanism to block a potential  $A \rightarrow C$  mapping. In none of these accounts is there any reason for an  $A \rightarrow C$  mapping ever to occur, for two sub-reasons:
- i. The process would only ever apply to final vowels, and only ever once
  - ii. The notion of Moreton's  $A \rightarrow B \rightarrow C$  schematic is incoherent, as no crucial reference is ever made to the  $V_1$  of a  $V_1V_2$  sequence.

I have discussed points (28a) and (28bii) at length throughout this section, but (2bi) has been less of a concern. However, it will become very important in the next section of the chapter, so I discuss it in more detail now.

If we are to understand the process in Hidatsa as being a regular morphologically triggered process of deletion, we must try and treat the morphology that occurs in imperative formation in Hidatsa as similarly as we can to other morphological processes, as in Bye & Svenonius (2012) and Zimmermann (2014). Processes like this have, in essence, two parts. The first part is the morphology, which in the simplest terms possible, can be described as any process that demands exponence. In effect, inflected forms should, all else being equal, be pronounced differently to uninflected forms<sup>27</sup>.

---

<sup>27</sup> Though, of course, zero-exponence is possible and not infrequent (see, for example, Trommer (2012) for a recent review). Also, I wish to be clear that I am not talking about multiple exponence, in terms of multiple affixes expressing one piece of morphological content. Whether multiple exponence can or should be admitted into a theory of morphology is a thorny question (it is banned under most interpretations of DM, see e.g., Halle & Marantz (1993), but allowed under word-based morphology approaches (see, e.g., A. Harris (2009))

There is no reason to suppose that morphology will apply more than once. Poser (1992) dates the concept of blocking, in which the application of a morphological rule precludes the application of any further morphological rules that have the same function, back to Paul (1896). It is perhaps most famously expressed in terms of the Elsewhere Condition (see Kiparsky (1973b)), where the application of a more specific morphological rule blocks the application of a more general rule. For example, the English regular plural {-z} does not apply to the word *child*, whose plural form is the irregular *children* (vs. \**childs*, \**childrens*, or \**childsren*).

Accounts of Hidatsa which take into account its essentially morphological character work in a way that can be argued to be directly equivalent. In both Bye & Svenonius (2012) and Zimmermann (2014), an affix of some sort is added to the final mora of a Hidatsa word. The subsequent deletion is a direct (if somewhat counterintuitive) result of this addition. The combination of the abstract feature and final vowel is uninterpretable, and the way that the constraints are set up in the language mean that deletion is the optimal way of resolving this tension. Thus, the morphology operates once and once only, and the phonology ensures that the eventual surface form is a licit one.

To conclude, as far as I know, there does not appear to be a class of processes that operates above the level of the segment that can be said to require any of the chain shift treatments that have been proposed and that I discuss in chapter 2. I have not discussed all such processes here, but, given the diversity of processes above the level of the segment (see the final section of chapter 4 for the details of this), I do not think that there is substantive evidence for a coherent class of processes at this level. For further discussion of deletion processes and stress-epenthesis effects (the two largest groups of processes above the level of the segment in the corpus) I direct the reader to Neasom (2011).

I believe that, at the very least, scholars who are interested in synchronic chain shifting should direct their attention to shifts that occur at the level of the segment. This is not a new position; as I have stated, it is one shared by, for instance, Baković (2013) and McCarthy (2007). However, I believe the contribution of this section is to

show the level of chaos that we introduce when we bring chain shifting above the level of the segment. We can see this purely from comparing the three shifts discussed in the previous subsections. The shifts very clearly have different motivations, they involve different processes, and they take place over different domains. An attempt to deal with all of these kinds of processes using the same method like Local Conjunction leads to representations that ignore the substantive differences between the processes and, at times, lead to incorrect predictions.

#### **5.4 Chain shifting at the level of the segment: Chain shift as smaller parts of something bigger**

In this section, I primarily explore two kinds of putative chain shift that are the most commonly cited in the corpus in chapter 4; consonant mutations and vowel height shifts. I use these as my test cases partially because of their frequency in the corpus. Unlike the shifts in the previous section, which had very different motivations, reflexes, and effects, these two kinds of shifts do appear to have genuinely unified purposes. This means that explanations applied to a particular example of these kinds of process should, in principle, be able to be applied more generally. Also, these effects are well-attested, and present in multiple language families, meaning that the worries about data that I discussed in section 5.2 are less pressing.

It is fair to ask why, at this point, I do not address tone sandhi effects in this section of the thesis, given that they are also one of the significant groupings of chain shift effects that we see in the corpus. A key reason for this is the high level of variation that we see between individual tone sandhi processes. When we consider consonantal mutation processes, there are several trajectories that the mutations can take; however, most of them are at least sonority-increasing on traditional views of the sonority hierarchy. There is also evidence that some consonant mutations can be sonority increasing (see Jensen (1994) on SeSotho, for example). What does not appear to be possible is a consonant mutation paradigm with several steps in which some are sonority increasing and others are sonority decreasing<sup>28</sup>.

---

<sup>28</sup> There are cases in Southern Bantu languages where certain mutation mappings are sonority decreasing, but the pattern as a whole is not. For example, Zoll (1995) discusses a pattern in Luganda where all consonants spirantize before reflexes of historical ‘super-high’ back vowels, realized in the synchronic grammar as a class of the high back vowel [u]. Underlying /r/ and /l/ are realized as [v] on the surface, a sonority increasing mapping. However, this is perhaps to be expected, given that *all*

Similarly, in the corpus there are examples of vowel height processes where height is incrementally increased or incrementally decreased, but not examples of processes where some of the mappings are increases in height and others are decreases. In all cases, the movement is unidirectional. This is not the case with the tone shifts in the corpus. Consider the  $HM \rightarrow \uparrow H \rightarrow H$  shift in Shuijingping Hmong, in which the  $A \rightarrow B$  part of the shift involves raising and the  $B \rightarrow C$  part involves lowering. The other shift in this dialect, schematized  $LML \rightarrow MH \rightarrow ML$ , is even less amenable to a straightforward analysis, as the second two tones in the sequence are less complex than the first. It is not even the case that tone sandhi shifts all have this ‘up-then-down’ profile. Consider, for example, the  $H \rightarrow M \rightarrow L$  shift in A-Hmao, which has a consistent downward trajectory, or the circular tone shift in Xiamen. Whilst there are a great many interesting issues surrounding tone sandhi shifts, they would necessitate tangential discussions that are beyond the scope of this chapter. I acknowledge that this would be an invaluable topic for future research, but for the moment I leave this issue aside.

Unlike in previous sections, in which I addressed individual chain shifts and discussed broader issues through the prism of the isolated effect, in this section I discuss the overall classes of mutation and vowel shift processes. I begin by explaining why I wish to treat the two classes of processes together. I then note that it is important to consider that the chain shifting part of both of these sets of processes sometimes constitutes only a sub-feature of the overall effect. I provide examples from within individual languages and from cross-dialectal comparison that suggests that chain shifting is an accidental by-product of the operation of these processes.

This lends weight to the idea that the  $A \rightarrow B \rightarrow C$  orders that we observe are, as in the Hidatsa effect discussed above, sometimes what we would expect from the basic principles of morphology operating once and once only, and the phonology mediating the response to this operation. This suggests that approaches to the

---

consonants are realized as either [v] or [f] in this morphological context, depending on whether they are voiced or voiceless underlyingly. This is a case where there is always one mutation mapping which is indifferent to the change in sonority it causes. Therefore, it does not represent a process where there are multiple mappings giving rise to a sonority decreasing trajectory.

problems posed by mutation and metaphony should not need to invoke any special mechanism in the grammar to block  $A \rightarrow C$  mappings. I argue that the fact that sometimes, particularly in OT, such a device *does* seem to be necessary is not a knock-down argument in favour of including such a mechanism in the grammar. It may, in fact, be a suggestion that there is a problem with the architecture of OT, or with the kinds of representations that we are using. Most radically, it may be necessary to take these kinds of chain shifts *outside* of the phonological grammar. I discuss pre-existing work in which all of these approaches are discussed.

By illustrating the epiphenomenal nature of chain shifting at the level of the segment, and showing that there are approaches to it that do not *require* the explicit blocking of an  $A \rightarrow C$  mapping, I argue that approaches that do require such a blocking do not offer us any real insight into why  $A \rightarrow B \rightarrow C$  mapping processes should occur. Rather, they offer us a false insight: that the phonology is overly concerned with attaining specific outputs. Instead, I argue, the morphology is concerned purely and simply with exponence, and the phonology is concerned only with creating *licit* outputs. Whether these form a chain shift or not is accidental.

#### 5.4.1 Mutation and metaphony – rationale for treating them together

At first blush, it may seem like it is a case of overgeneralization to bundle together consonant mutation and shifts in vowel height. In this section I argue both that there are certain properties that are shared by both kinds of process, and also that these properties are germane to my overall discussion in the same way. For concreteness, I first present data from a metaphonic vowel shift. Rather than Lena Spanish, I present data from the Servigliano dialect of Italian (data reproduced from chapter 4, from Kaze (1989, p. 39)):

(29)	(a)	modɛst-a	‘modest’ (f. sg.)	modest-u	‘modest’(m. sg.)
	(b)	mɔr-e	‘he dies’	mor-i	‘you die’
	(c)	kred-o	‘I believe’	krid-i	‘you believe’
	(d)	fjɔr-e	‘flower’	fjur-i	‘flowers’

The data shows an analogous pattern to the Lena Spanish data we have seen so many times. In certain morphological contexts, low-mid vowels raise to become high-mid

vowels (29a-b). In similar contexts, high-mid vowels raise to become high vowels (29c-d). In this dialect, as in Lena, all of the morphological endings that trigger metaphonic raising are high vowels, either /i/ or /u/. Importantly, we will see later on in this section that this is not always the case. For now, it is enough to know the basic facts that are true of all apparent chain-shifting metaphony processes: 1) The process involves vowels gradually changing in their specification for height; 2) this gradual change in height *definitely* has a morphological trigger; 3) this gradual change in height *may* have a phonological trigger (as in Lena and Servigliano Italian); 4) this phonological trigger is *not* suggestive of a general, genuine phonological rule operating independently of the morphology in the synchronic grammar. To my knowledge, no one has argued that there is an unconditioned, synchronic rule of metaphony that works like genuine vowel harmony in any dialect of Italian. This is particularly striking, given that there are some dialects that involve the interaction between metaphony processes and more general, phonologically motivated vowel harmony effects (for a modern example of an analysis in which this explicit division is made, see Mascaró (2011) or Mascaró (2015) on Servigliano Italian).

Bearing in mind these four qualities, I present data below from a well-studied consonant mutation shift in Irish, which is commonly known as eclipsis (see, e.g., Massam (1983), Ní Chiosáin (1991), Rice (1993), Grijzenhout (1995), Gnanadesikan (1997), Trommer (2009) for various theoretical approaches). The data below comes from Gnanadesikan (1997, pp.96-97), citing Ní Chiosáin (1991). Forms with /ʲ/ are described as ‘slender’, or palatalized, as opposed to forms without which are described as ‘broad’.

Table 5-10: Radical and eclipsed forms of consonants in Irish

Radical →	Eclipsed	Radical →	Eclipsed	No Change
p, p'	b, b'	b, b'	m, m'	m, m'
t, t'	d, d'	d, d'	n, n'	n, n'
k, k'	g, g'	g, g'	ŋ, ŋ'	h
f, f'	v, v'			s, s', l, l', r, r'

(30)	RADICAL	ECLIPSED
(a)	<b>t'</b> ax 'a house'	s'axt <b>d'</b> ax 'seven houses'
(b)	<b>k</b> ootə 'a coat'	ə <b>g</b> ootə 'their coat'
(c)	<b>g'</b> atə 'a gate'	wur <b>ŋ'</b> atə 'your (p.) gate'
(d)	<b>d</b> orəs 'a door'	ə <b>n</b> orəs 'their door'
(e)	<b>b</b> oskə 'a box'	ər ə <b>m</b> oskə 'on the box'

In eclipsis mappings, voiceless consonants are voiced (30a-b), whilst voiced consonants are nasalized (30c-e). The contexts for eclipsis are not phonological and are, to my knowledge, always described in morphosyntactic terms. Let us now consider the main salient generalizations present in this kind of consonant mutation: 1) the process involves consonants gradually changing in their specification for manner and voice; 2) this step-wise change in manner *definitely* has a morphological trigger; 3) this step-wise change in manner *definitely* does *not* have a phonological trigger; 4) the fact that there is no phonological trigger is indicative of the fact that there is no general, genuine phonetic process of voicing and nasalization occurring in the phonology of Irish.

It is worth saying something more about point 1) here, given that it may seem like there is a less obvious unitary characterization in the overall mutation process available than there is in a metaphony account. I would argue against this position in two ways. The first is that in terms of function the process seems to be a unified effect. As in metaphony, the same set of morphological triggers is the cause of both parts of the shift. As well as this, some theorists have pointed out that whilst the eclipsis mappings do not form what we would consider to be a typical lenition trajectory, for example (though there is another kind of mutation in Irish, called

lenition, that will be of interest later), one can discuss the movement as a unified increase in sonority.

Therefore, on some theories (see, especially, Trommer (2009)) it is explicitly assumed that the process operates on a unified scale of sonority. This assumption rests on the notion that sonority is purely a matter of voicing. If, as Harris (2006) points out, sonority is to be considered as an analogue of *intensity*, then it is unclear whether voiced or voiceless stops are more sonorous, given that the release phase of voiceless stops is significantly more intense than the release phase of voiced stops. Gnanadesikan (1997) labels the scale she uses to model the eclipsis effect the ‘Inherent Voicing’ scale, which is perhaps more accurate than using the term ‘sonority’.

On the flipside of this argument, whilst it is often assumed that vowel raising processes like metaphony can and should be couched as a gradual increase in vowel height (see, e.g., Maiden (1991), Parkinson (1996), Kirchner (1996), Gnanadesikan (1997)), it is not entirely clear that the first step in a shift like the Servigliano effect above is genuinely a shift in height. For example, a study by Grimaldi et al (2010) - cited by Calabrese (2011) in his discussion of approaches to metaphony - presents ultrasound evidence that the difference between /ɛ/ and /e/ in the Tricase dialect of Southern Italian is genuinely one of tongue advancement, rather than an increase in vowel height. Under those auspices, most metaphonic analyses lose their unified nature to some degree. Whilst it may be true that, all else being equal, the two vowel movements /ɛ/ → [e] and /e/ → [i] both lead to an increase in F1, so there is a unified phonetic parameter to consider, this suggests that (at least in this dialect), there is a difference in the articulatory process behind each of the movements. It should be noted that some previous accounts of vowel raising processes that use standard articulatory features, like Kirchner (1996), attempt to square this circle by talking about the phonetically unified nature of the process whilst using the features [+high] and [+/-ATR].

The overall point to be made here is that both metaphonic raising and eclipsis mutations can be seen as unified on some phonetic dimension, but it is hard for either to be seen as unified in terms of phonological representations, which

potentially accounts for why both mutation and metaphony are so resistant to straightforward phonological analysis. In sum, I believe that we can fruitfully consider cases of mutation and metaphony together in this section. I believe that their similarities, in that they are both morphologically triggered processes that affect particular natural classes of sounds in ways that people have suggested are unified, outweigh their differences (e.g., some metaphonic processes have phonological triggers, metaphony acts on vowels whilst mutation effects are usually observed on consonants).

I make it clear now that, perhaps disappointingly, this does not mean that I propose to finish this section with a new, perfect, unified solution for metaphony and mutation processes. This is not a goal that fits within the scope of this thesis, and it may not even be a goal that is genuinely desirable. Solutions that can model both of these processes already do exist. It would be possible to provide a Local Conjunction analysis of both kinds of process, for instance. As stated above, my argument is not that this does not work, but that it is not an insightful way to discuss how these kinds of processes work. The reason for my conflation of mutation and metaphony in this section is simply that most of the points that I wish to make about these kinds of processes will apply with equal force to both kinds of process, as well as similar vowel shifts in other language families. It is partially, then, in the interest of not repeating myself that I analyse the two kinds of processes together.

#### *5.4.2 Chain Shift as a smaller part of something bigger*

Placing an explicit mechanism into the grammar to stop  $A \rightarrow C$  mappings in cases of vowel shift or mutation suggests two things:

- 1) All else being equal, an  $A \rightarrow C$  mapping is what would result from the operation of these processes
- 2) This is a bad thing; for some phonologically important reason, these mappings must be precluded

I dispute both of these points, and in this sub-section make the argument that when general patterns of mutation and metaphony are considered, it is unimportant whether chain shifts result from the application of these kinds of morphology. I do

not attempt theoretical analysis of the patterns that I present at this point, given that my reasoning for showing them is primarily to illustrate the diversity of such processes.

#### 5.4.2.1 Lenition mutation in Irish

As I mentioned in the previous subsection, there are two separate kinds of initial consonant mutation in Modern Irish. In eclipsis mutation, we do observe a fairly regular chain shift pattern (/p t k/ → [b d g], /b d g/ → [m n ŋ]), but this regularity is nowhere near as apparent in the lenition mutation. The data below comes again from Gnanadesikan, who it should be noted does explicitly discuss both eclipsis and lenition mutations as chain shifts (forms and data from Gnanadesikan (1997)):

Table 5-11: List of lenition alternations

Radical →	Lenited	Radical →	Lenited	No Change
p, p'	f, f'	f, f'	∅, ∅'	
t, t'	h, h'	s, s'	h, h'	h
k, k'	x, x'			
b, b'	w/v, v'	m, m'	w/v, v'	
d, d'	ɣ, j, ɣ'	(N, N')	(n, n')	
g, g'	ɣ, j, ɣ'	(L, L', R, R')	(l, l', r, r')	

(31)	RADICAL		LENITED	
(a)	<b>p</b> o:stə	‘married’	in’ <b>f</b> o:stə	‘marriageable’
(b)	<b>t</b> ’ax	‘a house’	mə <b>h</b> ’ax	‘my house’
(c)	<b>k</b> ’ark	‘a hen’	mə <b>x</b> ’ark	‘my hen’
(d)	<b>f</b> ’i:əkəl	‘a tooth’	mə i:əkəl	‘my tooth’
(e)	<b>s</b> ’o:l	‘sail’	<b>h</b> ’o:l	‘sailed’
(f)	<b>d</b> ’e:ntə	‘done’	do <b>j</b> e:ntə	‘impossible’
(g)	<b>g</b> ’n’e:həx	‘faceted’	il’ <b>ɣ</b> ’n’e:həx	‘multifaceted’
(h)	<b>b</b> ’aləx	‘a way’	mo:rv’aləx	‘a highway’

There are, ignoring the difference between broad and slender consonants, but including the fact that initial /h/ is realized faithfully, 13 separate mappings present in Irish lenition. Of these, just two form a chain shift, /p/ → [f] and /f/ → [∅].

This can be argued to be largely a product of the inventory. /k/ lenites to [x] for instance, and by analogy we might expect initial /x/ to lenite to either [h] or [∅]. However, Irish does not have phonemic initial /x/, so this is not a question that it is possible to answer. This, then, does not constitute direct evidence against the idea that the effect is not a coherent chain shift overall, merely that the process is structure preserving and sometimes the inventory precludes chain shifts. A more interesting point in the data is presented in (31b). Whilst initial /p/ and initial /k/ lenite to their fricative counterparts [v] and [x], /t/ lenites to [h], neutralizing with instances of both underlying /s/ and instances of underlying /h/, which does not lenite. This neutralization is particularly troubling for accounts where chain shifting is part of the phonological grammar, because it is an instance where mutation leads *precisely* to an A → C mapping.

Gnanadesikan (1997) and Trommer (2009), who cites Pullman (2004), give the same reason for this. Whilst /t/ → [s] would be a structure preserving move, it would mark a change in stridency, either from [-strident] to [+strident], or ∅ to [strident]. Why it is so important to avoid a change in the specification of stridency is not discussed by either of the authors, and the spectre of circular reasoning hovers over this argument.

Q: Why does /t/ change to [h] and not to [s]?

A: Because of a highly ranked constraint ruling out stridents

Q: What is the justification for this highly ranked constraint on stridents?

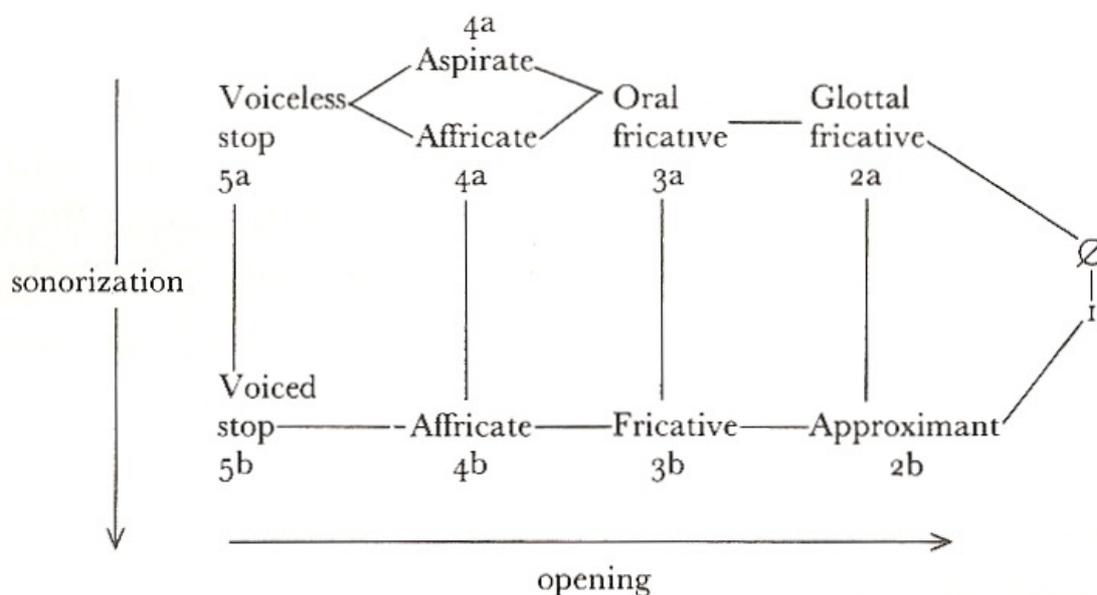
A: Because /t/ changes to [h] and not [s]

A potential basis for the t → h change is discussed by Kirchner in his thesis, where he asserts that “unaffricated stops never lenite to strident fricatives... such as [s] or [f]” (1998, p. 99). However, this is equally problematic. On Kirchner’s view, /f/ is a strident, which means that the Irish data, if genuinely synchronic, are counterexamples to his generalization, given that we observe a /p/ → [f] mapping. Additionally, alternations in English, such as *electri[k] → electri[s]ity* have been

argued to be productive (see, e.g., Pierrehumbert (2006)) and show just the kind of lenition that Kirchner argues to be impossible; an unaffricated obstruent leniting to a [+strident] sound<sup>29</sup>. Whilst it is easy enough to place a constraint blocking a /t/ → [s] mapping into the grammar (something like Gnanadesikan’s RESIST[*strident*], where segments must keep their underlying stridency), it is far from clear that this is a move that is justified by the data.

Secondly, it is worth pointing out that the one part of the shift that can be called a chain shift, /p/ → [f], /f/ → [∅] is not, strictly speaking, an A → B → C mapping. Consider a fairly standard view of lenition, as typified by Lass, from whom I take the diagram in 5-4 (1984, p. 178):

Figure 5-4: Possible lenition trajectories from Lass 1984



Irish does not have phonemic aspiration. As in English, stops are aspirated foot-initially, but this distribution is allophonic. Nor does Irish have phonemic affricates. Therefore, if Lenition mutation is a structure-preserving process, which it does seem to be throughout the paradigm, we can consider the /p/ → [f] mapping to be a genuinely one-step map (5a → 3a in the diagram). However, this does not explain why underlying /f/ should map to ∅, given that the glottal fricative /h/ not only exists in Irish, but is the target for lenition in the oral fricative /s/. Therefore, even in the

<sup>29</sup> I thank John Harris for bringing this to my attention.

mapping that appears to be a chain shift in Irish, what we appear to have is an  $A \rightarrow B \rightarrow D$  map, rather than an  $A \rightarrow B \rightarrow C$ . In terms of Lass' diagram, the last step of the shift is a mapping from 3a  $\rightarrow$  1.

Green (2006, p. 1961) also makes the point that it is difficult to see the rationale for the  $f \rightarrow \emptyset$  mapping under a purely phonological approach, in his discussion of the failure of Gnanadesikan's account to give a truly unified account of lenition mutation. He also makes the point that, on Gnanadesikan's account, more than one scale would need to be invoked to explain the full effect, given that she describes the alternations between /N, L/ and [n, l] as a "decrease in length or tension" (1997). This is a good point, but it does not necessarily speak against a chain shift reading per se, given that they do appear to be minimal movements.

Finally, it is also worth noting the mappings from broad /b/ to [w], and broad /d/ and /g/ to [j] (which according to Swingle (1993, p. 452) are regular mappings, not in free variation as is indicated in Gnanadesikan's table above) are also non-minimal changes. On Lass' numbering system, it is a move from 5b to 2b. We can see that the minimal change can be observed in Irish, because the minimal change is observed in the slender counterparts, in which, for example, /b'/ surfaces as [v'] and not [w']. Indirectly, we also know that broad [v] is a possible surface allophone, because it is what /f/ maps to under eclipsis mutation. In the case of Irish lenition mutation, then, we can see that whilst there are plenty of cases of the kind of minimal movement we associate with chain shifts, it is rare that a full,  $A \rightarrow B \rightarrow C$  chain shift is the result of this movement. Some of the reasons for this are straightforward. It is more puzzling, and more troubling for accounts where neutralizations are precluded purely by grammatical machinery that neutralizations (as in /t/  $\rightarrow$  [h], /s/  $\rightarrow$  [h]) and non-minimal movement (/f/  $\rightarrow$  [ $\emptyset$ ], /b/  $\rightarrow$  [w]) take place across the pattern.

#### 5.4.2.2 *Vowel lowering mutation in Thok Reel*

As mentioned in section 5.4.1, many metaphony shifts have what appear to be phonetic triggers. In Lena Spanish, for example, the suffixes that trigger metaphonic raising all contain the high vowel {-u}. This complicates the analysis of processes of this nature somewhat, as it is not clear whether the phonological trigger is in any way salient to the speaker. There are, however, vowel height effects in which it is very

clear that there is no phonological trigger that speakers would be able to recover. In this section, I discuss a language which can be said to contain processes that are, in essence, vowel mutation effects. The main discussion will focus on the Nilotic language Thok Reel, and is based on recent fieldwork by Reid (2010), making reference to the one theoretical treatment of the effect that seems to be available (Trommer, 2011).

Thok Reel has the seven-vowel system /i e ε a ɔ o u/. Each of these vowels has two variants, which Reid labels ‘breathy’ and ‘modal’. In a paper that comments extensively on Reid’s work, Trommer (2011) labels the two kinds of vowels ‘breathy’ and ‘creaky’. There is a persistent difference between the 2<sup>nd</sup> and 3<sup>rd</sup> person singular forms of verbs and all other conjugations in the Thok Reel present tense, as shown in table 5-12 (data from Reid 2010, p.75. Underlining denotes breathy voice).

*Table 5-12: Underlying vowels and their surface realizations in Thok Reel*

	1	2	3	4	5	6	7	8	9	10	11	12	13
2/3SG	<u>i</u>	e	<u>e</u>	ε	<u>ε</u>	a	<u>a</u>	ɔ	<u>ɔ</u>	o	<u>o</u>	u	<u>u</u>
Elsewhere	<u>jɛ</u>	ε	e	a	<u>a</u>	a	<u>a</u>	a	<u>a</u>	a/ɔ	o	wɔ	<u>wɔ</u>

Reid points out that the 2/3 singular form is the most likely to be the underlying form, even though there are more variants in the elsewhere condition in the Thok Reel present tense (1SG, 1PL(incl), 1PL(excl), 2PL, 3PL). This is because if the elsewhere form is assumed to be the underlying, there are three potential mappings for underlying /a/ ([a], [ε], [ɔ]), two for underlying /a/ ([a], [ɔ]), and, as can be seen in (10) in table 5-12, above, an underlying specification that exists in free variation between /a/ and /ɔ/. This offers convincing evidence that the 2/3SG vowels are the underlying forms, and in many cases these are unambiguously higher than their elsewhere forms. Reid discusses the effect as being a lowering process overall, but does not explicitly call this overarching process a chain shift.

Trommer (2011), however, is quite explicit that the effect constitutes what he terms a ‘vowel-lowering chain shift’. He presents a schematic that shows the chained nature of particular parts of the lowering process (p. 194):

(32)  $\underline{e} \rightarrow e, e \rightarrow \varepsilon, \varepsilon \rightarrow a$

Examples of these particular shifts are shown below (Reid, 2010, p. 78):

(33)  $\underline{e} \rightarrow e$

		‘awaken (someone)’	‘prevent’
(a)	3SG	k <u>é</u> eer	p <u>é</u> een
(b)	1PL	kéeer-kɔ(n)	péeen-kɔ(n)

$e \rightarrow \varepsilon$

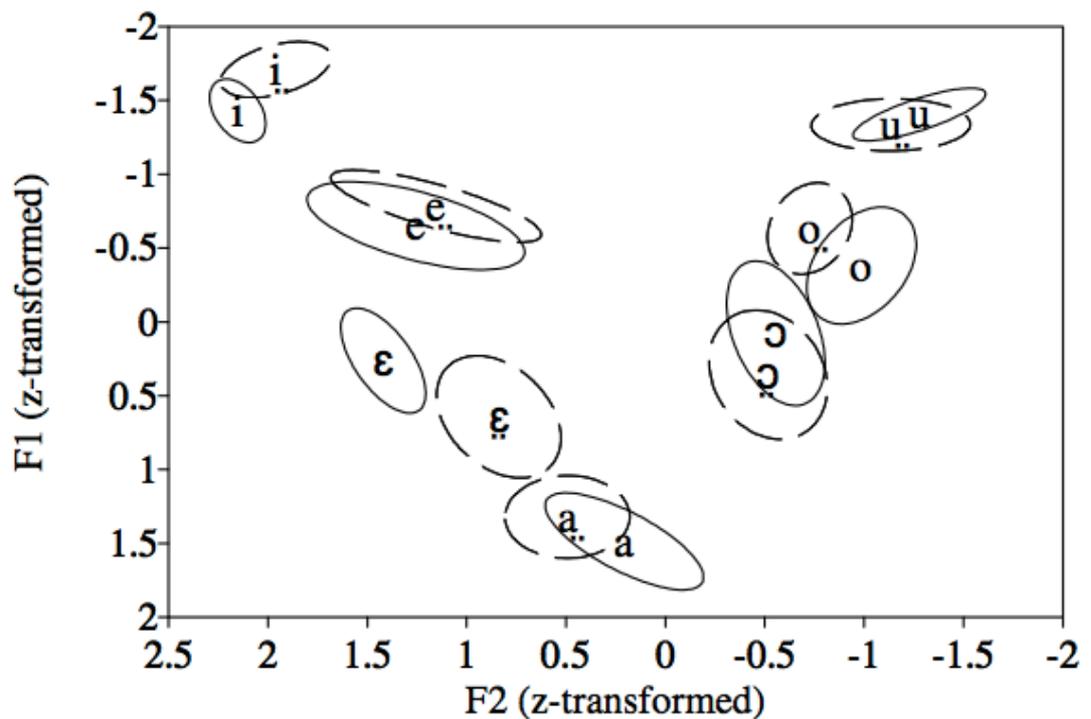
		‘dig’	‘count’
(c)	3SG	teT	kwéen
(d)	1PL	tɛt-kɔ(n)	kwɛn-kɔ(n)

$\varepsilon \rightarrow a$

		‘know’
(e)	3SG	ɲɛC
(f)	1PL	ɲaC-kɔ(n)

The first part of the shift ( $\underline{e}/ \rightarrow [e]$ ) is somewhat problematic if we are to maintain the general idea that the shift is a coherent lowering effect. In order to preserve a unified analysis in which each of the 13 mappings in the table above constitutes lowering (unless there is nowhere to lower to, as in 6 and 7), Reid and Trommer both assume that the breathy vowels are systematically higher than their modal counterparts. However, it is not clear from the measurements Reid takes that this is true.

Figure 5-5: Acoustic measurements of Thok Reel vowels from Reid 2010<sup>30</sup>



Whilst overall the values for breathy e and o are slightly higher than their modal counterparts, it does not appear to be by a significant amount. Indeed, when discussing the diagram as a whole, Reid notes that whilst the values for /o/ and /i/ do appear to be higher than their modal counterparts, for most of the vowels, the situation is much less clear: “We can see that with the rest of the vowels /e, a, ɔ, u/ there is *considerable overlap* in formant frequencies for breathy and modal phonemes” (p.64, emphasis mine). What is more, the value for breathy /ɛ/ is reliably *lower* than its modally voiced counterpart. Reid does not perform a statistical analysis of the average F1 distinctions between vowels, so it is impossible to know if any of the differences between breathy and voiced versions of vowels are genuinely significant. However, the fact remains that there are examples of breathy vowels being higher, lower, and indistinguishable from individual modal vowels in terms of height. This suggests that Trommer’s assertion that “creaky vowels have systematically lower F<sub>1</sub> than their breathy counterparts” (p.196) is too strong. In my

<sup>30</sup> Reid and Trommer mark breathiness with two dots under the vowels in question. I use underlining instead.

opinion, there is not enough evidence to support the labeling of breathy → modal alternations in the data as lowerings.

The main issue, though, is that there is no way in which the overall effect can realistically be seen as a coherent chain shift, a fact that Trommer tacitly acknowledges in his statement “the phonological changes are systematic but still inconsistent” (2011, p.194). There are several of these inconsistencies that I discuss in turn. The first and most obvious of these inconsistencies is the behavior of the high vowels, whether they are breathy or modal. In all cases, high vowels diphthongize, with the mappings /i/ → [jɛ], /u/ → [wɔ], u → [wɔ]<sup>31</sup>. We can see that the ‘chain shift’ mapping would be a possible realization in these cases (e.g. /i/ → [ɛ]) but instead the usual realization is the decidedly non-chain shift diphthongization. The chain shift mapping would be a ‘simpler’ realization than the diphthong in two ways.

Firstly, it would not introduce an additional segment that is not present in the underlying representation. Secondly, even if we wish to discard the fact that diphthongization has occurred and focus only on the quality of the vowel involved in the diphthong, we see that this is still not an example of chain shift style movement. If it were, we would expect that the diphthongal form for underlying /i/ would be either [je] or [jɛ], as this would count as a one-step movement if only vowel quality was considered. Indeed, in antipassive constructions, there appears to be optionality in this regard, at least with regard to underlying /u/, which can be realized as [wu] or [wo] (Reid, 2010, p. 80). Reid’s description suggests that this variation is lexically governed. In the general case, however, we have a movement that skips a potential target that would give us a chain shifting mapping, and this cannot be explained with simple reference to the inventory of the language.

This is not the only instance of this kind of skipping. Take, for instance, the mappings that Trommer describes explicitly as forming a chain shift: /ɛ/ → [e], /e/ → [ɛ], /ɛ/ → [a]. We see that the second two mappings are changes in phonological

---

<sup>31</sup> There are no verbs in Reid’s dataset with underlying breathy /i/. I concur with Trommer (2011, p.197) that if these verbs do exist we would expect their mutated forms to diphthongize to [jɛ], by analogy with the other high vowels.

height, at least on the assumptions made by Reid, who describes Thok Reel as a four-height system (2010, p. 54). This, on its own, is not necessarily troubling for a chain shift account. The assumption would be that the chain shift genuinely works on *height*, meaning that, in essence, you would expect two identical chain shifting patterns, one for the modal vowels and one for the breathy vowels. This is bolstered by the mappings we observe in the diphthongs, in which breathy underlying high vowels surface with their breathy specification intact, and the same is true for the modal vowels. This is cast into doubt, though, by the first mapping of the ‘shift’. This mapping on the other hand, does not involve a change of vowel height, but a change of voicing, from (in Trommer’s terms), breathy to modal voicing.

The logic of Trommer’s analysis, which I do not discuss in detail here, is that the shift is a generalized increase in sonority. He asserts that what corresponds to each level of height on Reid’s analysis can be further subdivided in two, with the breathy vowel on each level slightly more sonorous than the modal vowel. This suggests that genuinely *minimal*  $A \rightarrow B \rightarrow C$  movement should result in a chain shift of the form  $/\underline{e}/ \rightarrow [e]$ ,  $/e/ \rightarrow [\underline{e}]$ ,  $/\underline{\varepsilon}/ \rightarrow [\varepsilon]$ ,  $/\varepsilon/ \rightarrow [a]$ ,  $/\underline{a}/ \rightarrow [a]$ ,  $/a/ \rightarrow [a]$ . This is not a chain shift that we get, though we do get parts of it ( $/\underline{e}/ \rightarrow [e]$ ,  $/a/ \rightarrow [a]$ ). This puts any explanation of the Thok Reel effect that seeks to see genuine unity in trouble, at least in terms of mapping relations. If we are to define a minimal,  $A \rightarrow B$  mapping as a movement in sonority, then we are unable to explain not only the diphthongization that we observe, but also the fact that we do not get the minimal chain shift discussed above. If, on the other hand, we define a minimal,  $A \rightarrow B$  mapping as a one-step change in phonological height, ignoring voicing, we are still left with no easy explanation of the diphthongization, and we cannot explain why  $/\underline{e}/$  does not map to  $[\underline{e}]$ , or indeed why  $/\underline{o}/$  maps to  $[o]$  instead of  $[\underline{o}]$ .

Finally, it is worth considering the variable mapping that underlying modal  $/o/$  undergoes in Thok Reel. As both Reid and Trommer point out, when  $/o/$  is preceded by  $/w/$ ,  $/o/$  is realized as  $[\underline{o}]$ , which on an understanding of the effect in which a minimal movement in height constitutes a chain shift, would count as an  $A \rightarrow B$ , potentially chain shift mapping. However, the mapping that does not require any phonological conditioning (essentially the elsewhere mapping) is  $/o/ \rightarrow [a]$ . This, again, is exactly the kind of  $A \rightarrow C$ , neutralization mapping that we would not

predict to co-exist with chain shift mappings in a system where the phonology actively enforces a chain shift mapping. This is why it is so important that /o/ → [a] is the elsewhere mapping. If it were the case that ordinarily we would see /o/ → [ɔ], but in some tightly circumscribed phonological context, the neutralization mapping was forced, it would be a more justified move to introduce a minor rule or specific constraint simply to deal with the non-chain shift mapping. It appears, in this instance, that the reverse is true. If the morphological processes involved in exponing person features are operating normally, then we get the neutralizing map /o/ → [a]. It is only in a particular, purely phonological context that the chain shift mapping obtains.

The overriding message of the previous two subsections is that seemingly more ‘well-behaved’ chain shifts operating at the level of the segment can be difficult to explain in terms of the basic mapping relations that are central to most definitions of chain shifting. Within a particular language or dialect, it is possible for a mutation pattern affecting vowels or consonants to simultaneously display chain shift mappings, neutralization mappings, and completely unexpected mappings (like Thok Reel diphthongization) that do not seem to interact with the rest of the system at all.

#### *5.4.2.3 Chain shift metaphony as an epiphenomenon of rule-based approaches*

In the previous two subsections, we saw that there are processes of vowel and consonant mutations in which apparent chain shifts coexist with decidedly non-chain shift mappings, suggesting that a grammatical explanation that is based around blocking *any* particular mapping operation is not a valid way of dealing with such processes. However, it is clear that there are processes in which the mapping relations are not so complex. This is why Lena Spanish has been our guiding example throughout this thesis. Every vowel that can increase in height by one step does so. As a first, general point, it is worth noting that the possibilities for the operation of metaphony in Lena are limited by two forces. The first is that Lena Spanish has the five-vowel inventory /i e a o u/. The second is that, in standard Lena Spanish, the process is structure preserving. If we take Maiden’s (1991) general approach to historical metaphony, that it constitutes gradual phonetic raising towards post-tonic high vowels, then the movements we observe will only ever be an increase in vowel height. With this in mind, a chain shifting pattern is really the only way that

the effect in Lena could possibly go, apart from neutralization. Similar arguments obtain for dialects like Servigliano Italian.

Again, taking a historical perspective on metaphony suggests that we should not be entirely surprised by the fact that we end up with dialects like Lena Spanish, or Servigliano Italian. The account that Maiden gives of historical metaphony is predicated on the idea that “metaphony became established first in the (more susceptible) high mid vowels, secondly in the low mid vowels, and finally in /a/” (1991, p. 125). The simplest possible way that we could look at this would be to conceptualize the two parts of the change in a chain shift dialect as linear rules. As Maiden’s book focuses on Italian dialects, and as it is his discussion I am following, I use Servigliano, as opposed to Lena, for my examples. In standard feature terms, the rules that we would need are below:

- (34) (a) /-high, -low, +ATR/ → [+high] / \_\_{i, u}  
(b) /-high, -low, -ATR/ → [+ATR] / \_\_{i, u}

The first rule ensures that high-mid vowels are realized as high vowels in a variety of morphological contexts (I use {i, u} as a shorthand for these contexts). The second rule states that low-mid vowels are realized as high-mid vowels. If we follow a strand of thought in rule-based phonology that has been present since Kiparsky (1973a), and is perhaps most explicitly outlined in section 3 of Bromberger & Halle (1989), we can suggest that the rules are bound to appear in the order above, because the order of rules in a phonological grammar reflects the order in which they enter that grammar. This ordering relation is, on this account, a fact of history, rather than an extrinsic solution to a particular problem. It is worth noting that this is an analysis that can easily be extended to other dialects of Italian, which do not show a chain shift mapping. I give data from some such dialects below.

(35) *Foggia: Low mid vowels raise to become high (Valente (1975), Calabrese (2011))*

(a)	péte	piti	‘foot-SG/PL’
(b)	kjéna	kjínu	‘full (f sg)’/’full (m sg)’
(c)	grós:a	grús:u	‘big (f sg)’/’big (m sg)’
(d)	móf:a	múf:u	‘soft (f sg)’/’soft (m sg)’

(36) *Grado: No alternations in low-mid vowels (Walker (2005), Calabrese (2011))*

(a)	bélo	béli	‘beautiful-SG/PL (m)’
(b)	béla	béle	‘beautiful-SG/PL (f)’
(c)	mét-e	mít-i	‘he/she/it put’/’you put’
(d)	mórto	mórti	‘dead-SG/PL (m)’
(e)	mórta	mórte	‘dead-SG/PL (f)’
(f)	fjor	fjur-i	‘flower-SG/PL’

(37) *Calvello: Diphthongization to glide/high-mid diphthongs in low-mid vowels (Gioscio (1985), Calabrese (2011))*

(a)	pére	pjéri	‘foot-SG/PL’
(b)	péntsa	pjéntsi	‘he/she/it feels’/’you feel’
(c)	mése	mísi	‘month-SG/PL’
(d)	fórte	fwórti	‘strong-SG/PL’
(e)	móve	mwóvi	‘he/she/it moves’/’you move’
(f)	kavróne	kavrúni	‘charcoal-SG/PL’

(38) *Francavilla Fontana: Diphthongization to glide/low-mid-front diphthongs in low-mid vowels (Sluyters 1988, p.164, citing Rohlfs 1956-61)*

(a)	lnta	ljéntu	‘slow (f sg)’ / ‘slow (m sg)’
(b)	frédda	fríddu	‘cold (f sg)’ / ‘cold (m sg)’
(c)	gróssa	grwéssu	‘big (f sg)’ / ‘big (m sg)’
(d)	pilósa	pilúsu	‘hairy (f sg)’ / ‘hairy (m sg)’

Considering the Foggia dialect first, there is complete neutralization of all non-low vowels to high. This could be treated with a feeding analysis, where the rule turning

low-mid vowels into high-mid preceding the rule that turns high-mid vowels into high vowels. However, if we take the approach that rule ordering reflects historical order, this would require a further rule change. This rule change would not be unnatural, indeed it is the kind of rule-change that Kiaprsky (1973a) predicts, because it leads to a transparent rule. However, classical accounts like these do not demand that this kind of reordering should occur, only that it may, and that this reordering may not necessarily be extrinsic.

Maiden does not argue for this kind of representation, but the logic of his general argument is that the reason for the variation between dialects that we see in (a-d) is that the general process of metaphonic raising was morphologized at different stages. If we make this assumption, coupled with the assumption that high-mid vowel raising *always* precedes mid-vowel raising, then rather than reordering, we can simply suggest that the rule affecting mid-vowel raising in Foggia is the neutralizing rule, below:

(39) /-high, -low, -ATR/ → [+high, +ATR] / \_\_{i, u}

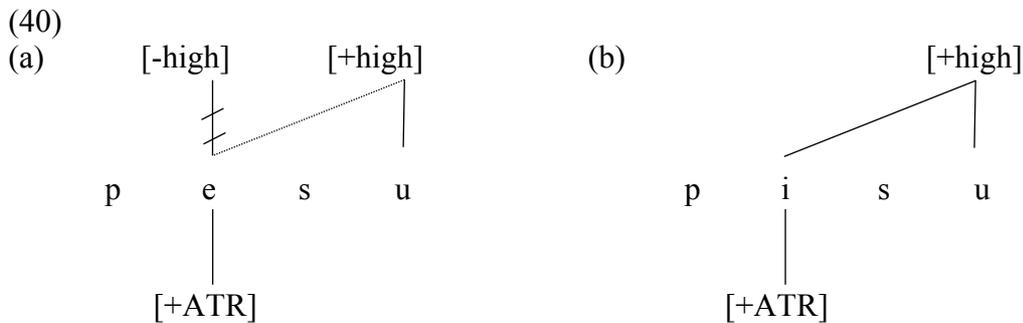
Diphthongizing dialects can also be treated in this way, with the high-mid vowel raising rule preceding the low-mid rule. This is actually a helpful move for a dialect like Calvello, in which the diphthongization produces high-mid vowels. If the ordering were reversed, the high-mid vowel forming the nucleus of the diphthong would be a potential target for the high-mid raising rule, which would lead to [ji] and [wu] diphthongs. It is worth pointing out that these would presumably be disallowed anyhow, as a general principle of Italian is that it does not allow diphthongs of this nature (see, e.g., Krämer (2009, p. 54)). However, the ordering of the high-mid rule before the diphthongization rule has the epiphenomenal benefit of stopping this problem from ever occurring.

In dialects like Grado, where no change occurs in the low-mid vowels at all, the simplest solution is that, again following a historical explanation, metaphony did not spread to the low-mid vowels in these kinds of dialects. Thus, there is no rule affecting low-mid vowels at all. The benefit of this kind of straightforward, rule based account is that it treats all dialects in essentially the same way. There is an

ordering that is imposed by the historical ordering of the language. Whether a chain shift is the result of the order is purely a function of when the low-mid vowel process became morphologized.

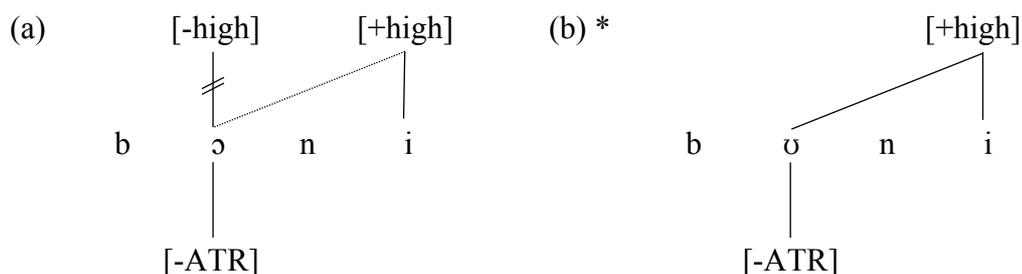
However, this is not a popular kind of analysis in present-day phonological theory. There are sociological reasons for this, the rise of Optimality Theory, which does not admit ordering of processes, being one. There are also more linguistically grounded objections (see Chapter 2, section 2 for discussion of this point). However, a more elaborate rule-based account, using autosegmental representations, sidesteps these objections but also does not work on the principle that the primary goal of chain shifting metaphony is to block an A → C mapping.

Calabrese’s approach to metaphony has been refined over various iterations (1985 ; 1995; 2005), and is one of the more widely cited approaches to the problem. Across all dialects, the basic process is the same; the metaphonic suffix spreads the feature [+ high] onto the preceding stressed vowel. If that stressed vowel is /e/ or /o/, both of which have the feature specification [-high, -low, +ATR] then this is unproblematic, as the feature set [+high, -low, +ATR] is allowable in all Italian dialects:

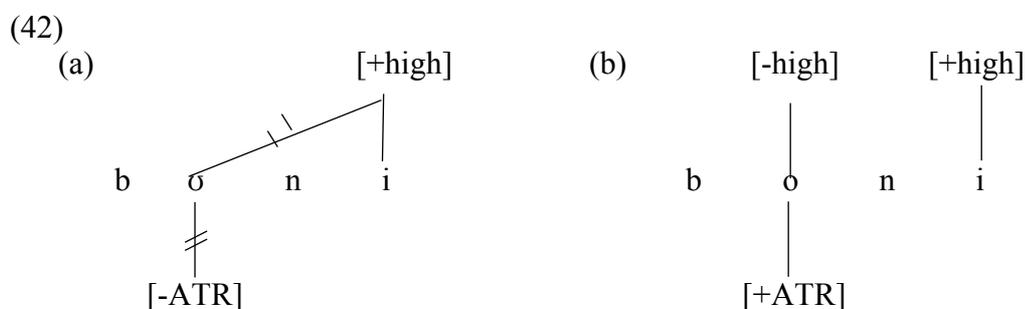


As can be seen in these heavily simplified representations, the feature [+high] spreads and delinks the original [-high] feature, thus /e/ surfaces as [i]. The process is less simple when the stressed vowel is [-ATR]:

(41)



The configuration created in the diagrams above is not allowable in Servigliano, by dint of structure preservation. Like many Italian dialects, Servigliano has the same 7-vowel system as the well-known Tuscan variety: /i e ε a ɔ o u/. Spreading [+high] onto a vowel whose underlying featural makeup is [-high, -low, -ATR], the resulting output is [ɔ], a [-ATR, +high] vowel. As this is not a phonemic vowel in Servigliano or other dialects featuring metaphony, Calabrese argues that there is a constraint of the form \*[+high, -ATR] in *all* such dialects. The consequences wrought by this constraint vary from dialect to dialect. In the cases that we are particularly interested in, the violation of the \*[+high, -ATR] constraint results in a reversal of the values of *both* of the relevant features, i.e. a specification of [-high, +ATR], a high-mid vowel:



In Calabrese's analysis, this operation (termed 'negation', (Calabrese (1995, p. 400), or 'excision' in Calabrese (2011)) takes the form of a complete delinking of the offending configuration. By default, the reverse specification for each feature for place is realized on the surface. The process by which we get neutralization is very similar, and is termed 'delinking' by Calabrese. In this case, the [+high] specification remains associated to the stressed vowel, with only the [-ATR] specification delinked. The resulting outputs are the [+ATR, +high] vowels [i] and [u]. This gives the neutralization pattern that is observed in the Foggia dialect. The motivation for

this (avoidance of [-ATR, +high] segments) is exactly the same as the motivation for the chain shift dialects, and the choice between ‘negation’, ‘delinking’, or the repair strategy that Calabrese claims causes diphthongization, which he terms ‘fission’, appears to be essentially parametric. It is beyond the scope of this chapter to do justice to Calabrese’s account of metaphony, and I give no space at all to competing rule-based accounts (those who are interested can consult, for example, Kaze (1989), Maiden (1991), Puente (1996), Nibert (1998), or Cole (1998)). However, there are two key issues that are raised by Calabrese’s account in terms of the characterization of Servigliano-style metaphony as a chain shift<sup>32</sup>.

Firstly, as shown above, Calabrese introduces a constraint that prevents certain mapping relations from occurring. This appears to be familiar from other chain shift theories, which employ constraints with two specific functions: initiating the shift by militating against underlying /A/ surfacing faithfully; or blocking of a potential A → C mapping. However, this constraint does not block either of those mappings. Instead, it blocks a mapping on the far more routine grounds of structure preservation. It is a constraint on the inventory of the language, not on the kinds of processes that the language can instantiate. There is no special reason why a chain shift is preferred to neutralization.

Indeed, we can see that both patterns exist, and it is only a parametric choice of repair strategy that gives us one or the other. The corollary of this is that on Calabrese’s analysis, the lack of an A → C mapping is again unimportant, and in a crosslinguistic sense, the chain shift is an epiphenomenon of a constraint on structure preservation that is uninterested in the mapping relation. Essentially, as long as surface [ɪ] and [ʊ] are disallowed, the realization of lax mid vowels is unimportant. Sometimes a chain shift will be the result of the repair of this constraint, sometimes it will not. Calabrese is explicit that we should expect that the process should be essentially the same across dialects, which accords with the closely related history of dialects featuring metaphony. Calabrese sums up his analysis thus: “[T]he range of variation should be limited to the range of results produced by the simplification procedures, which is what we find” (1995, p. 400).

---

<sup>32</sup> I also note that, as far as I am aware, Calabrese never refers to this kind of metaphony as a chain shift. For an account that does, see Cole (1998).

With this in mind, it is possible to address the second key issue, which is representational. Consider how Calabrese’s negation analysis works. First, [+high] spreads across, creating the disallowed structure [+high, -ATR]. Then, and only then, can the excision occur. The repair must occur *after* the spreading of [+high]. This is of crucial importance, because the typical conception of chain shifts in rule-based phonology is that they are counterfeeding processes (see above, and the discussion in chapter 2). The B → C part of the shift occurs first, followed by the A → B part of the shift. This is why they are problematic for rule-based accounts, because they require the extrinsic ordering of rules.

A further problem, as noted by, for example, Łubowicz (2011), is that extrinsically ordered rules are incapable of expressing functional unity if the two rules are instantiations of the same overall process. Both of these problems are solved at a stroke if we use a Calabrese-like analysis of metaphony. Functional unity is hard-coded into this kind of analysis. The purpose of metaphony is always to spread [+high]. This is true whether the system under consideration is the five-vowel system of Lena (as discussed in Hualde (1989)) or the seven-vowel system of Servigliano. In terms of ordering, the derivation in table 5-13 shows that in a bizarre unattested system where the rules were reversed, we would get incorrect outputs:

Table 5-13: Derivation for the incorrect, counterfeeding order

	/bɔn+i/	/pes+u/
Negation	n.a	n.a
[+high] spreading	bɔni	pisu
	*[bɔni]	[pisu]

If, however, the excision repair is *fed* by the spreading of [+high] (which is also demanded by the basic logic of the situation), then the putative chain shift results:

Table 5-14: Derivation for the correct, feeding order

	/bɔn+i/	/pes+u/
[+high] spreading	bɔni	pisu
Negation	boni	n.a
	[boni]	[pisu]

This, then, is an example of a seeming chain shift that can be analyzed on a rule-based system featuring only standard articulatory features, the basic operations of spreading and delinking, and intrinsically ordered, transparent rule-interactions. It takes place in order to satisfy the well-attested linguistic principle of structure preservation, and the fact that a chain shift order results seems to be totally epiphenomenal. In essence, if structure preservation is to be obeyed, and the only features being manipulated are [+high] and [+/-ATR], then there are three potential monosegmental mappings that can be created (I ignore diphthongization here. However, as Calabrese points out, both *parts* of the diphthong, for example the [j] and the [e] of [je], are phonemes of the relevant dialects). These are [ɛ, ɔ], [e, o], and [i, u]. On the assumption that teleological concerns such as the amount of contrasts present in the system or amount of deviation from the underlying form are important, we would not expect to see systems in which all three of these mappings were present.

If, on the other hand, we make the assumption that the choice is essentially parametric, we would expect that some system would employ each of the three mappings, with scant regard for such concerns. The latter case appears to be closer to the truth. In rule-based phonology, then, we have a straightforward way of modeling the  $A \rightarrow B \rightarrow C$  order present in metaphony. However, things are not so simple when we consider OT analyses.

#### *5.4.2.4 OT accounts of metaphony – Chain shifting has to be special, and that is a problem*

There are various Optimality Theoretic approaches to metaphony, some specific to particular dialects (see, e.g., Mascaró (2011) on Servigliano or Gnanadesikan (1997) or Finley (2009) on Lena), and others that aim to cover more ground, attempting to present a typology (a well-known example is Walker (2005)). In this section I discuss a recent attempt to give an OT typology of Italian metaphony effects by Mascaró (2015). I note that the version of this article that I refer to is the longer version available on Mascaró's website, rather than the cut-down version that appears in Torres-Tamarit, Linke & van Oostendorp (2016). I choose to study Mascaró's approach to Italian metaphony partially because of its currency, and partially because it takes an explicitly typological approach to the problem at hand.

Like the previous section, Mascaró considers dialects in which metaphony takes the form of neutralization, diphthongization, chain shifting, and regular shift of high-mid vowels with no concomitant shifting in the low-mid vowels.

As well as this, Mascaró makes an explicit distinction between morphologically triggered processes like metaphony and general phonological processes of vowel harmony. This did not really come up in the previous subsection, because in a sense it was not especially important for the discussion of Calabrese’s theory. Indeed, if we look back at the roots of autosegmental theory, the distinction between phonology and morphology is not judged to be especially important (see for example Lieber’s assertion that “autosegmental phonology and morphology are not distinct theories” (1987, p. 5)). However, this is a more pressing concern in OT.

Following Finley (2009), Mascaró discusses processes like metaphony as instances of ‘morphemic harmony’. Essentially this means that metaphony is not conditioned by the high features of the suffix vowel *per se*, but rather by floating features whose function is purely morphological. Finley gives two arguments that suggest that a morphemic harmony treatment is a more sensible analysis even of a language like Lena Spanish, which is her test case. Firstly, a morphemic harmony treatment allows for a uniform analysis of metaphony dialects that contain suffixal high vowels as well as dialects where these triggers do not exist. For example, in the Sonnino dialect of Italian (discussed in Savoia (2015, p. 249)), we can see the same kind of chain shift raising pattern that we observe in Servigliano Italian, but without overt triggers causing these vowel changes:

(43)	(a)	pete	pete	‘foot-SG/PL’
	(b)	prete	prete	‘priest-SG/PL’
	(c)	novo	novo	‘new-Fem.SG/Masc.SG’
	(d)	nove	nove	‘new-Fem.PL/Masc.PL’
	(e)	mese	mise	‘month-SG/PL’
	(f)	nofe	nufe	‘walnut-SG/PL’

Secondly, as I have already stated, metaphony processes are not usually envisioned as general harmony processes. In general, metaphony applies only to the tonic

vowel, even going so far as to skip post-tonic vowels in languages like Lena Spanish (though see, for example, Mahanta (2012) on a variety of Bengali in which a similar effect causes raising in multiple vowels).

Mascaró's specific approach to metaphony interprets this morphemic harmony as a set of features that form a floating suffix, specified as [+high, +ATR]. To ensure that this floating feature is realized, and to make sure that it is realized on the stressed vowel, Mascaró uses FLOAT-STRESS a constraint taken from Wolf (2007b). This constraint assigns a violation to any form in which a floating feature is not linked to the stressed vowel of a word. In the case of Mascaró's [+high, +ATR] floating suffix, any form which fails to associate either feature to the stressed vowel violates FLOAT-STRESS twice, and any form in which one floating feature is attached but the other is not violates FLOAT-STRESS once. As well as a constraint mandating that the floating features be associated with the stressed vowel, Mascaró also uses an altered version of Łubowicz's preservation of contrast constraints, in order to prevent neutralization.

Recall that in Łubowicz's preservation of contrast analysis it is not individual inputs being evaluated, but rather scenarios, or sets of outputs. Mascaró attempts to bring the principles of contrast preservation into a more standard OT architecture. Although he is not the first to do this (see, for example Flemming (1995)), it should be noted that his use of Contrast Preservation constraints introduces a third kind of constraint into the grammar, that does not work in the same way as standard markedness or faithfulness constraints.

The question then becomes how necessary it is to introduce a constraint that is this complex. I argue that this device is completely unnecessary in non-chain shift dialects. This is important, because the fewer contexts in which contrast preservation is useful, the less justification there is for entirely changing the nature of the grammar to accommodate it. In appendix A, I show in detail that, even under all of the other assumptions of Mascaró's analysis, the preserve contrast constraint is unnecessary in all dialects except for the chain shift dialect. It is possible to simply remove the constraint from the analysis for dialects featuring neutralization. This makes sense, given that contrast is *not* preserved in these cases. Additionally, the

constraint can simply be removed from dialects where only high-mid vowels raise and low-mid vowels are realized faithfully. To model diphthongization dialects is slightly more complex without the Preserve Contrast constraint, and requires some slight constraint re-ranking, but I argue that it is, in essence, possible. However, it is not possible to re-order the constraints in a way that will model a chain shift dialect if the Preserve Contrast constraint is removed, as table 5-15 shows.

Table 5-15: Chain shifting dialects under Mascaró's analysis (2015, p.15)

/ε/	DEP-C	PRCONT(STRESS, MID)	FLOAT-STRESS	ID(ATR)	ID(hi)
ε			**!		
↻ e			*	*	
i		*!		*	*
jε	*!				

It appears that the PRCONT constraint is necessary in Servigliano-like dialects. However, rather than this being a boon for theories like Mascaró's, it is actually a problem, given that it is *only* in the chain shifting dialect that Contrast Preservation appears to be a necessity.

On rule-based accounts of metaphony, there is no great mystery to why A fails to map to C. If we take the historical viewpoint, the blocking is an unavoidable consequence of the low-mid vowel rule being added to the grammar after the mid-vowel raising rule. If we take on a Calabrese-like analysis, *all* metaphony mappings occur for the same reason (avoidance of lax high vowels) and a chain shift order is just one of a set of parametric repair responses. On these kinds of analysis, the lack of an A → C mapping is just as lacking in mystery as the examples in previous sections. It is only on an OT account that we need special machinery to model chain shifts. If we consider the range of metaphony effects, that special machinery is needed only to model the chain shift dialects. This seems problematic, given the closely shared history of dialects featuring metaphonic raising, and the general principles of OT.

Consider Pater's (1999) well known study of \*NT effects in Austronesian languages. Pater shows that cross-linguistic reactions to a similar problem can be modeled effectively through the permutation of the same set of constraints. It is a necessary corollary of this method of analysis that in some languages or dialects, individual constraints will not be of any great importance. However, there is a difference between demoting, let us say, IDENT(high) to a level where it has no great effect in a particular dialect, and introducing a constraint that is not only totally unnecessary in every dialect except one of the dialects that you have under consideration, but must be ranked ahead of *any* other constraint.

If a constraint must either hold sway over everything, or else is completely irrelevant to the process, then it is no more than a *deus ex machina*. In a case where we have every reason to believe that the variation we encounter across closely related dialects comes from the same source, and shows no signs cross-dialectally in being especially interested in preserving contrast, or stopping certain degrees of movement down a scale, chain shift cases present a problem that none of the other patterns do for Optimality Theoretic analyses. Put simply, why do Servigliano speakers care about preserving contrast, or not moving too far down some phonetic scale, or not violating multiple constraints, when that is not a concern for speakers of Grado, Foggia, or Francavilla Fontana?

A potential reason for this is OT's insistence on teleological explanation. Assume that historical metaphony, à la Maiden, was a gradual, phonetic assimilatory process that affected mid-high vowels and then low-mid vowels. Once morphologized, synchronic metaphony is no longer necessarily an assimilatory process, and in fact *cannot* now be an assimilatory process in some dialects, like Sonnino, as there is nothing to assimilate to. Whilst much has been made of the fact that OT is good at dealing with situations in which multiple repairs are observed for the same problem cross-linguistically (e.g., Prince & Smolensky (1993/2008), Pater (1999)), the architecture requires that there is a specific, productive, synchronic reason why repair strategy 'x' should be used in language 'y'. Put simply, these repair strategies must be expressible in terms of the constraint set of a given language.

The problem with this is that it is completely ahistorical. If we assume that Maiden is correct in his assertion that the reason that we observe the kind of cross-dialectal variation that we do is because of the stage at which metaphony was morphologized, this provides us with an explanation for the variation. A learner of the dialect will hear the forms of their dialect and form a grammatical interpretation of how the forms are made. If we make the classical OT assumption that all speakers have the same toolkit from which to build their grammar, speakers in chain shift dialects have a problem. This is because they will need a completely different type of tool (contrast preservation, or constraint conjunction, or scales) to speakers of other dialects.

#### 5.4.2.5 Representations

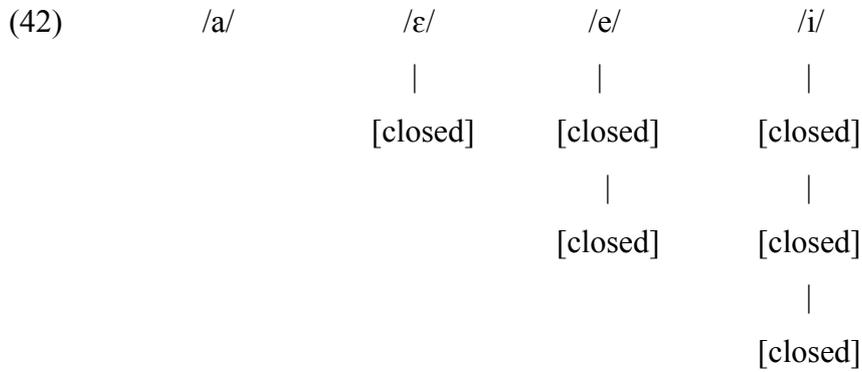
Perhaps, though, this is being unfair. Perhaps it is not OT that is at fault, but rather the system of representations that is used in most OT and rule-based accounts, binary articulatory features. As Maiden points out, “generalized vertical adjustments in vowel height are notoriously the graveyard of binary features” (1991, p. 136). There are many approaches to vowel height that attempt to change the nature of vowel height features in order to capture the kinds of processes that we observe (Puente (1996) contains a good review of the pre-OT literature on this).

Parkinson (1996) and Nibert (1998) both discuss metaphony processes in terms of additive height features (both based on the work of Clements (1991)). In essence, these approaches state that rather than features such as [+/- high], vowel height is instead expressed by numerous instances of a feature [open] (Clements (1991), Nibert (1998)) or [closed] (Parkinson, 1996), each of which appears on a different node in an autosegmental representation:

*Parkinson’s height feature table for Servigliano Italian (1996, p. 38)*

	/i, u/	/e, o/	/ɛ ɔ/	/a/
[closed]	*	*	*	
[closed]	*	*		
[closed]	*			

This effectively means that the representation of /ɛ/ contains one instance of [closed], whilst, for example, the representation of /i/ contains three (diagrams adapted from Parkinson (1996, p. 39)):



Under this kind of account, a chain shift is totally unmysterious. The process is simply the addition of an instance of [closed]. Parkinson uses both autosegmental representations and Optimality Theoretic constraints to model this effect.

Table 5-17: Parkinson's analysis of Servigliano (1996, p.40)

/mɔri/	MAX[cl]	ALIGN[cl]	IDENT[cl]
☞ mor i \   \   [cl]			*
mɔr i     [cl]		*!	
mɔr e <[cl]>	*!		
mur i \   \   \ [cl] \   \   [cl]			**!

The fully faithful candidate [m̥ori] is ruled out by the constraint ALIGN[cl], which mandates that [cl] be associated with the stressed vowel (for the full formulation, see Parkinson (1996, pp. 38-39)). The neutralization candidate [muri] incurs more violations of IDENT[cl] than the chain shift candidate IDENT[cl]. Under this approach, chain shifting can be explained as a simple resolution to a simple conflict; ALIGN[cl] mandates that some linkage of the feature [closed] occur. The passive phonology of the language, represented simply by the IDENT[cl] constraint, mandates that the satisfaction of ALIGN[cl] is achieved by the simplest means possible. Chain shifts in this kind of framework are combinations of a constraint forcing some kind of change, and the basic units of OT, faithfulness constraints, restricting the nature of that change.

This interaction of basic faithfulness and a constraint mandating that some change should happen has also been applied to consonant mutation effects in Irish. Trommer (2009) suggests that lenition is caused by an affix that works to increase the degree of sonority present in the initial consonant, and that a highly-ranked constraint is required to ensure that the morpheme is realized on the surface (Trommer uses REALIZE MORPHEME (RM) (see, for example, van Oostendorp (2005)). However, the nature of the affix that Trommer suggests, like Parkinson's [closed] feature, deviates from standard features. Trommer, following de Lacy (2002), suggests, in a way that is not dissimilar to Gnanadesikan, that features like sonority are essentially scalar. These scales are defined by a system of grid-marks, as in table 5-18 (from Trommer (2009), slide 15):

*Table 5-18: Trommer's sonority scale*

	Representation	Abbreviation
Voiceless obstruents and laryngeals	son: oo	son:
Fricatives and approximants	son: xo	son: x <sub>1</sub>
Laryngeals (and ∅)	son: xx	son: x <sub>2</sub>

The difference in sonority between stops and fricatives is represented by the addition of a grid-mark (x). This representation allows Trommer to shift some of the burden

of the analysis of lenition in Irish from computation to representation. Rather than stating that the M morpheme must cause *complete* change in inherent voicing, and leaving it entirely to the constraints to block full movement down the scale, Trommer suggests that the sonority morpheme always attempts to do something else; if possible, add a grid-mark.

$$(43) \bullet \leftarrow \rightarrow x_{\text{son}} \text{ (slide 22)}$$

In the standard case, then, /p/ is a labial stop. The addition of a sonority grid-mark causes a change from /p/ → [f]. Place, voicing, etc. all stay the same. The tableaux in table 5-19, which I have adapted slightly, illustrate Trommer’s analysis.

Table 5-19: Trommer’s lenition analyses

/x <sub>son</sub> + p/	RM	ID[cont]	ID[nas]	ID[voi]
p [son: ]	*!			
f [son: ]	*!	*		*
v [son: x <sub>1</sub> ]		*!		*
☞ b [son: x <sub>1</sub> ]				*
m [son: x <sub>2</sub> ]			*!	*

/x <sub>son</sub> + b/	RM	ID[cont]	ID[nas]	ID[voi]
b [son: ]	*!			
v [son: x <sub>1</sub> ]	*!	*		
☞ m [son: x <sub>2</sub> ]			*	

In the first tableau, Realize Morpheme rules out the fully faithful candidate [p], as well as the fricative [f]. This is because the morpheme can *only* be realized by the addition of a grid-mark. The voiced fricative [v] does not violate RM, as it is a realization in which a grid-mark has been added. However, it does violate ID[cont], because the input /p/ is a stop. This leaves a decision between [b] and [m]. Again, the choice is made by a basic faithfulness constraint, in this case ID[nasal]. [b] is preferred over [m], even though both of them realize the sonority morpheme, because it satisfies more (and higher ranked) basic faithfulness constraints than [m].

In the second tableau, the fully faithful candidate [b] is again ruled out by Realize Morpheme, as is its fricative counterpart [v]. It is worth pointing out that Parkinson's and Trommer's analyses still do rely on what is essentially the notion of a scale, something that I have argued against in previous sections (see especially chapter 2). Whilst it is true that on both accounts, vowel height and sonority respectively are considered to be properties that have multiple levels, the only computation that is involved in each tableaux above is a simple process of addition. In simple terms, the morpheme (addition of a [closed] feature or a grid-mark) is completely blind to what it is being added *to*. If a [closed] feature or sonority grid mark can be added, and if this addition creates a licit phonological structure, then all to the good. If not, then it may be the case that the closest alternative segment in terms of adherence to general faithfulness results in the addition of multiple features or gridmarks (a two-feature change) or the lack of addition of a feature or gridmark (no change, and a failure to align the feature or realize the morpheme).

Neither of these processes involve counting, or constraints that refer to the scale. Indeed, Trommer does not believe that regular sonority affixation can ever result in the addition of more than one grid-mark. Crucially, the sonority grid system does not involve the grammar being burdened with the knowledge of notions like numbered values, like Gnanadesikan's ternary scales, or adjacency, like her IDENT-ADJ constraint. The addition of the sonority gridmark, or the recasting of height features as additive [closed] can thus be argued to be smaller departures from standard phonological theory than the recasting of contrasts in voicing as a scalar feature.

However, it is clearly not a situation that we wish to find ourselves in that we have to change the system of representations that we use simply in order to deal with particular, recalcitrant effects (see section 5.2.2 for a similar argument). Trommer's account is explicit that sonority gridmarks are independent of phonological features and can be used to model consonantal and vocalic processes. The same sonority gridmark account is used by Trommer to model both vowel mutation in Thok Reel and consonant mutation in Irish. However, this is an admission that the phonological grammar must be able to attend to a property that *cannot* be represented using standard features, meaning that there are essentially two sub-systems of

representation in Trommer's grammar; one of abstract features, and one based on sonority that more directly corresponds to the phonetic reality of the language. Whether one finds this problematic or not, it is worth noting, as I have previously in section 5.4.2.2, that Trommer's sonority scale of Thok Reel does not seem to exactly correspond with the phonetic reality in that language. If the sonority feature is not directly correlated with genuine phonetic sonority, it is hard to see how it operates.

In sum, if we wish to analyse vowel height or consonant manner chain shifts using OT, it seems that we need to accept that either we need to change fundamental assumptions of how OT works, or abandon the assumption of standard features. Both of these seem to be fairly drastic measures, although I would argue that, counter-intuitive though it may seem, the recasting of height features is a less damaging move than the introduction of an  $A \rightarrow C$  blocking mechanism, given that the data presented throughout this chapter gives us little reason to suppose that the motivation behind chain shifts is an explicit move by the grammar to block such mappings. The overall picture that emerges, when one surveys the literature on metaphonic raising, or consonant mutation, is a series of fixes to the phonology, in which well-established principles appear to be bent to the will of a process which, I would argue, is fairly marginal in the world's languages.

There is, of course, another possibility. It may be the case that these kinds of putative shifts may not be part of the phonological grammar at all. In the following subsection, I briefly consider this possibility, and discuss accounts that have taken this position.

#### *5.4.2.6 Mutation and metaphony as morphology*

In previous sections I have discussed the idea that it is most sensible to treat metaphony and mutation as processes that are morphosyntactically triggered. This is the standard position for mutation, and has plenty of support in accounts of metaphony too (see Finley (2009), Mascaró (2015)). However, in most of these accounts, there is an interaction between the morphology, which introduces the change in the consonant or vowel, and the phonology, which mediates the change. This seems to be the prevailing method of dealing with at least metaphony. The approaches to mutation that I have discussed to this point are similar, in that the

phonology is necessary to determine what underlying forms should map to when they undergo mutation.

This is, perhaps, not the simplest way of addressing this problem. It is also possible that the entirety of mutation and metaphony processes are dealt with by the morphology, and that the phonology is bypassed entirely. The most hardline view that I have found on this topic is laid out by Kaye & Pöchtrager (2014) who state that the phonological grammar can only be involved in processes that meet four separate principles:

- (44) (a) The non-arbitrariness principle
- (b) The minimality hypothesis
- (c) The principle of semantic transparency
- (d) Structural limitations on phonological events

It is (44a) and (44b) which concern us here. The non-arbitrariness principle, which has its roots in Kaye, Lowenstamm & Vergnaud (1990, p. 195), essentially states that there must be a clear phonological motivation for a process to apply that can be read off from its environment. This would seem to immediately discount processes like consonant mutation, which are morphosyntactically conditioned, and dialects of metaphony where there is no triggering high vowel. The minimality hypothesis is summed up by the authors as “[n]o exceptions, no rule ordering, no non-derived environment effects” (2014, slide 4). On this kind of analysis, they note, no kind of metaphony is possible (slide 21). They cite Walker’s analysis of Grado and Veneto Italian, stating that the patterns are too complex and have too many exceptions to be viewed as a unified phonological process. The presentation does not suggest what *is* happening in these cases, though I note that this is not the authors’ intention. The idea of the presentation is simply to show that metaphony (and a range of other umlaut processes) do not exist as synchronic, phonological processes. There are, however, proposals where more detail is given to the processes that occur. I give one example below.

Green (2006) makes the proposal that all Celtic mutations are purely morphological, and do not involve any kind of phonological derivation. Instead, the radical and

mutated forms are both contained in the lexicon, and a morphosyntactic process, similar to case assignment, selects the correct allomorph in the correct context. The reason behind this lies in Green's attempt to hold fast to the principles of classical Optimality Theory. That is to say, on Green's view, only standard markedness and faithfulness constraints should be able to play a role in phonological computation. This is, in a way, a similar staking out of territory to that made by Kaye & Pöchtrager (2014). The only difference is that Kaye & Pöchtrager are defending the core tenets of Government Phonology (or at least, their version of the theory, known as GP 2.0 (Kaye & Pöchtrager (2013))), whilst Green is defending the core tenets of OT. In both cases, processes that do not conform to these principles must be placed somewhere else.

The problem for mutations, vis-à-vis the basic principles of OT, is that "universal markedness principles appear to play no role" in the process (Green, 2006, p. 1946). For a simple example, consider that, in their radical contexts, initial voiceless plosives are completely unobjectionable in Irish, but they are disallowed in mutation contexts. This cannot be explained by a highly ranked constraint against these segments. This, again, is a surprisingly similar objection to that raised by Kaye & Pöchtrager, in their invocation of the non-arbitrariness principle. Both principles essentially state that the phonology can only apply when there is a clear, phonological reason.

Green's analysis of mutation is reasonably straightforward. Potential targets of mutation are marked with a diacritic, a move seen as necessary by Green because of the lexical exceptions that abound in mutation processes (see Green (2006), section 3.4). Forms that can undergo mutation have two allomorphs in the lexicon. Green then represents the mutation agreement via an OT constraint which he calls MUTAGREE. MUTAGREE is violated by either the surface realization of a radical form in the presence of a mutation trigger, or the surface realization of mutated form *without* the presence of a mutation trigger. This strikes me as somewhat odd. Recall that the main thrust of Green's paper is that there is nothing phonological about mutation processes. It is, therefore, germane to ask why allomorph selection is enforced by an OT constraint that coexists in tableaux with purely phonological

constraints<sup>33</sup>. This strangeness is reflected in the fact that, at least in Green's analysis, MUTAGREE is inviolable (it is at least, always either the most highly ranked constraint in his tableaux or one of multiple, non-crucially ranked, highest ranked constraints). The fact that MUTAGREE must always be observed places it outside of the purview of traditional OT constraints.

Perhaps anticipating this objection, Green shows that the same generalization can be made using statements about the nature of the lexicon, in the style typical of word-based approaches to morphology (Green cites Ford et al. (1997) and Singh (1987) as the inspirations behind these kinds of representation). Without wishing to go into too much detail on this point, I show one such statement that Green proposes for the changing of stop consonants to their fricative counterparts in Irish lenition.

(45) “/C<sub>[-cont{labial, dorsal}]</sub>X/<sub>radical</sub> ↔ /C<sub>[+cont{labial, dorsal}]</sub>X/<sub>lenition</sub>” (p.1981)

Essentially, this states that any word (represented by the X) that has a labial or dorsal stop in its radical form will have a labial or dorsal fricative in its lenited form. As Green points out, this is not a rule. Rather, it is a bidirectional statement, giving preference to neither the radical nor the lenited form.

The point of this discussion is that, under this analysis, or under Pöchtrager & Kaye's view of phonology, the reading of these kinds of effects as chain shifts is totally incoherent. This is because there is no derivation at all taking place, but rather allomorph selection. The selection process is blissfully unaware of the effect that it has on the system. There is no need to explain the lack of A → C mappings, or the fact that sometimes we get A → B, A → C, or A → D mappings in the same language. Returning to the logic that I laid out in my discussion of Hidatsa, my baseline assumption is that the morphology is blind to non-morphological concerns like the inventory of phonological systems, or degree of movement on some phonetic scale. It simply acts to expone.

---

<sup>33</sup> I thank Florian Breit for making me think about this point, and for many illuminating discussions about mutation processes. Of course, all mistakes here are nothing to do with him.

These accounts, if true, spell real trouble for a chain shift analysis. However, approaches to problems like this run the gamut from purely phonological (e.g., Gnanadesikan's treatment of Lena Spanish) through a series of interactions between morphology and phonology (e.g., Trommer's accounts of consonant and vowel mutations in Celtic languages (2009) and Thok Reel (2011) respectively), to purely morphological accounts like Green (2006). This suggests that there is still a great deal up for grabs in the analysis of processes like these. A tentative attempt to address the empirical question of whether a metaphony like effect appears to be amenable to synchronic phonological analysis is presented in the next chapter.

### 5.5 Conclusion

In this chapter, I do not claim to have conclusively proved that there is no such thing as a synchronic chain shift. There are many shifts from the corpus that I have not discussed in detail here, mainly for reasons of space. I give a brief example below of another kind of process that has been cited as a chain shift effect, and my reasons for disagreeing that this is an accurate characterization. Again, these reasons centre around the lack of a consistent, genuine  $A \rightarrow B \rightarrow C$  mapping.

There are two effects in the corpus (found in Bedouin Hijazi Arabic and Palauan), which can be classed as vowel reduction shifts. Their schematics are shown in (46).

- (46) (a) Bedouin Hijazi Arabic:  $a \rightarrow i \rightarrow \emptyset$   
 (b) Palauan:  $u: \rightarrow u \rightarrow \text{ə}/\emptyset$

In some ways, these appear to be good potential examples of synchronic chain shift. The argument (put forward by Kirchner (1995; 1996) for BHA, and Zuraw (2002) for Palauan) that these processes represent unified effects of vowel reduction seems fair, which implies that there could be a genuine  $A \rightarrow B \rightarrow C$  mapping at work. However, there is room for concern even in the chain shift accounts of these processes. Firstly, Kirchner, citing Al-Mozainy (1981), points out that it is not necessarily the case that /a/ will reduce to an [i] that has the same phonetic character as underlying /i/:

“In non-final open syllables, short /a/ raises to a high vowel (transcribed [i], realized as [i], [u], or [ɪ], depending on adjacent consonants” (1996, p.341)

This means that we only observe a genuine  $A \rightarrow B$  mapping in a subset of cases, and whether or not we observe this mapping is less to do with a failure of neutralization than the phonotactic constraints of BHA. Similarly, in Palauan, the processes in the chain shift do not seem to be especially regular, as the schematic shows. There is variation between whether short vowels in Palauan are realized as schwa, or whether they are deleted entirely. Zuraw describes the variation thus: “Although consonantal environment is not a completely accurate predictor, in most cases it seems to determine whether consonant deletion occurs” (2001, p.3).

In both of these cases, I see no reason to disagree with Kirchner and Zuraw’s accounts in terms of the nature of the overall effect. However, it appears to again be a set of processes where the  $A \rightarrow B \rightarrow C$  mapping relation seems to be afforded more importance than is perhaps necessary. I would tentatively suggest that there is no a priori reason why vowel reduction should necessarily cause low vowels (in the BHA case) or long vowels (in the Palauan case) to syncopate completely. Indeed, this assumption would suggest that total vowel neutralization is what we should always expect in cases of vowel reduction, which is not the case (see, e.g., J. Harris 2005 for a selection of vowel reduction processes). Instead, an alternative hypothesis is that vowel reduction simply causes all vowels to reduce to a certain extent. The nature of this extent may be controlled, as both Zuraw and Kirchner suggest, by factors like surrounding consonants, but also by factors like frequency or speech rate. To make a judgement on the degree to which the putative shift is in any way categorical in either of these languages, it would be necessary to have access to high-quality data from speakers. Collecting and assessing data of this kind is a potential route that further research could take, that could erode support for the necessity of the grammar to intercede in seeming chain shift effects still further.

I have argued both that I do not believe that many of the effects that are present in Moreton’s corpus should be described as chain shifts, and that it is not necessary to include a specific grammatical mechanism blocking  $A \rightarrow C$  mappings in the grammar. In fact, I argue that this is a damaging move. One monolithic analysis of chain shift processes (i.e., something blocks the  $A \rightarrow C$  mapping) assumes that

substantive similarities exist between processes that have different functions and motivations. At the same time, because they operate at such a low level of detail, mechanisms of this nature may obscure genuine similarities between sets of processes, like mutation and metaphony.

It is obviously true that there are plenty of processes in the phonologies of the world's languages that contain  $A \rightarrow B \rightarrow C$  mappings. However, I argue that this is not, in and of itself, especially interesting. Given the variety of processes that exhibit this mapping, and the fact that there is no satisfying 'one-size-fits-all' explanation for these processes, the term 'synchronic chain shift' appears to be no more than a vague, descriptive label, rather than a problem that needs solving.

# Chapter 6: The learnability of synchronic chain shifts: Directions for empirical research and a pilot study

## 6.1 Introduction

Until this point, the thesis has been concerned with what a chain shift is, how the term should be defined, and how shifts should be modelled. This, of course, rests on the assumption that there is genuinely something to model. The final section of the previous chapter raised the spectre of a more radical possibility: that putative chain shifts of the kind that it is hardest to explain away (morphologically triggered vowel raising or consonant mutation effects) are not actually processes that involve the phonological grammar at all. Instead, they are straightforward cases of allomorphy, and the relevant variants have to be stored. If the variants are lexically listed, then there is no sense in which they can genuinely be thought to be chain shifts. The  $A \rightarrow B \rightarrow C$  mapping that lies at the bedrock of any existing definition of chain shifting would not exist, as the mapping is essentially shorthand for a derivation, in which underlying  $A$  is changed *by rule* into an instance of surface  $B$ , and so on and so forth.

Choosing between the two competing ways of looking at effects of this nature is not straightforward, and the choice is likely to be influenced by one's theoretical background. To illustrate this point, consider again the difference between the treatment of synchronic chain shifts in Rule-Based Phonology, Optimality Theory and Government Phonology. Chain shifting is an expected outcome in Rule-Based Phonology, as under such accounts, rules apply in sequence and can apply in multiple orders. In Optimality Theory, opacity is seen as a fundamental problem that requires solving if we are to say very much of substance about phonology (see especially the work of Baković (2007; 2013)). Chain shifts are a vital test case in this, as they are a seemingly common instantiation of underapplication opacity. By contrast, in Government Phonology, there is no account of chain shifting because it is not predicted by the theory. GP is an essentially declarative model of phonology; because of this, derivation with multiple steps is impossible. If there is no derivation, then there is no way of defining an  $A \rightarrow B \rightarrow C$  chain.

The trouble with all this is that there is little in the way of direct evidence that might favour any one of these approaches to chain shifting over another. In essence, a theorist's opinion on chain shift is directly connected to whether the theory they subscribe to predicts that chain shifts will exist. However, it is clear that the question of whether chain shifts do, in fact, require phonological computation is an empirical one. Indeed, there are already studies that have made attempts to answer the question, testing whether speakers of Bengali (Nagle, 2013; 2016) and Xiamen (Zhang, Lai & Sailor 2006) are able to generalize chain shift effects in their own languages to wug-words that contain the morphological triggers required to instantiate the chain shift.

In this chapter, I show how we might address this question from a different experimental perspective. I present the results of a pilot study, using the Artificial Grammar Learning (AGL) paradigm, that investigates whether there is a learning bias towards (or away from) vowel chain shifts when they are compared to similar, but minimally different patterns of vowel shift. In this study, three artificial language fragments were created, all identical except for certain morphologically triggered vowel shifts, all manipulating vowel height. The phonological complexity of these vowel shifts was manipulated in three distinct ways:

- (1) (a) featural complexity of the pattern (i.e., how many features does a learner need to change in order to accurately learn the entire pattern?)  
(b) coherence of the pattern (how many natural classes of sounds are involved in the pattern overall?)  
(c) transparency of the pattern (viewed in derivational terms, is the overall pattern transparent or opaque?).

The basic premise behind the experiments is this: If participants are using their phonology in *any* way, then we should see some kind of learnability advantage for at least one of the patterns over at least one of the others. The two non-chain shift patterns were chosen because whilst they differ only minimally in terms of the alternations that are involved, they differ significantly in their level of phonological complexity. Because it is still not clear which elements of phonological complexity are most important for learnability, any result in which the differences between

conditions is significant would be interesting for its own sake, particularly in the context of the AGL literature as a whole. The following sections give more information about the creation of the languages and the reasons for the specific methodologies and manipulations employed.

It is vital to point out two things at the outset of this chapter. The first is that the results of the experiment do not present good evidence that there is any one of the three patterns that is significantly more or less learnable than the others. In effect, what is presented here is a set of results that is largely negative. What positive results there are do not point to an overall difference in learnability, and are potentially effects of the task rather than genuinely phonological effects. With this in mind, it becomes important to reaffirm that the experiment set out to look for a positive result, i.e., that at least one of the conditions should be significantly different from at least one other condition in terms of the overall learnability of the pattern. Any significant differences between any of the two conditions are interesting for their own sake, but of primary importance is the notion that any such orderings strongly suggest that people use their phonology when learning chain-shift-like patterns. The reverse is not so clear. A negative result in the experiment certainly does not give us clear evidence that learners do *not* use their phonology in acquiring chain shift like patterns, as there are many possible reasons for a seemingly flat result. However, there are still many reasons to report and analyse these findings. I present three here.

*(i) Negative results should be published*

For some time now, scientists have been concerned with the ‘file-drawer’ problem (a term popularized in Rosenthal (1979)), the notion that there is a bias in the sciences towards publishing only significant results. Publishing negative results is a vital part of scientific enquiry for two separate reasons. The first is that there may well be methodological or other practical errors to be found in a given study. If the negative result is published, it allows for easy replication of the study with the errors removed. There are methodological errors in this pilot that I acknowledge at the outset may be obscuring significant differences between the conditions. It is important to report these and be clear about them as an advertisement of the pitfalls of these kinds of experiment. Secondly, it is important to consider that certain effects which are predicted to exist may well not exist. A negative result is potentially the

first step towards discovering that an effect or a phenomenon that has been assumed to have some kind of reality (either physical or psychological) is not, in fact, real.

*(ii) Negative results can tell you about the limitations of a paradigm*

AGL experiments have been part of linguistics for quite some time (e.g., Reber (1967) *et seq*) but it is fair to say that they are a particularly popular method of experimentation in phonology at present (see the review in Moreton & Pater (2012a; 2012b), for example). The study below introduces a rare level of complexity for an AGL task. Whilst there have been experiments that introduce opacity (e.g., Ettliger (2008), Ettliger, Bradlow & Wong (2014)) and patterns that involve multiple segments (e.g., Tessier (2012)), the majority of AGL experiments involve one phonological alternation (see, for example, Finley & Badecker's series of vowel harmony experiments, or White (2013; 2014) on phonological saltation). It is important to discuss the methodological issues that arise when attempting to teach more complex patterns.

*(iii) Maybe there will genuinely be no effect*

As alluded to in (i), simply because a phenomenon has been assumed to exist does not necessarily mean that it does. Previous studies on individual chain shifts have failed to find clear evidence that they are productive in wug-test scenarios (Zhang, Lai & Sailor (2006); Nagle (2013; 2016), See 6.3 for more discussion of this). It is, of course, notoriously difficult to prove a negative, but the more studies that fail to find a particular effect, the larger the weight of evidence against it. The crucial point in this is that if the negative evidence is simply taken to be a result of methodological or practical problems and ignored, it is not possible to build up this weight.

The following two subsections address overall methodological questions. I present justification for using vowel height shifts as the test cases for my experiment, and then justification for using the Artificial Grammar Learning paradigm. I then discuss the design of the experiment, before moving on to analyse the results.

## 6.2 Why vowel height?

As I have previously noted, a key reason that synchronic chain shifting is so difficult to define, and that different theories resist easy comparison with each other, is that each theory delimits the kind of effect that should be considered a chain shift differently. It is interesting then that vowel height shifts should be so widely attested, and so frequently invoked as a ‘canonical’ example of chain shifting (see chapter 4 for further discussion of this, and chapter 5 for examples of vowel height shifts that do not stand up to closer scrutiny). The three major doctoral dissertations treating synchronic chain shift in detail (Gnanadesikan (1997), Łubowicz (2003a), Mortensen (2006)) all address the issue of a vowel height chain shift at some stage in their analysis. Moreover, Moreton & Smolensky (2002), whilst providing their corpus that displays chain shifts from a wide range of domains, exemplify the concept with a raising shift from Western Basque, previously discussed in Kirchner (1995).

Articles which discuss specific instances of synchronic chain shift also often focus on vowel height. Kirchner (1996) (a reworked version of Kirchner (1995)), discusses a three-step raising process in Nzèbi ( $a \rightarrow \varepsilon \rightarrow e \rightarrow i$ ) using a Local Conjunction analysis allied to scalar constraints. Schmidt (1996) provides a detailed discussion of the Basaa shift mentioned above using a rule-based account hinging on underspecification. Nagle (2013), the first attempt at an empirical study of a vowel height shift, discusses an  $\text{æ} \rightarrow e \rightarrow i/\text{ɔ} \rightarrow o \rightarrow u$  process in Bengali which has also been extensively discussed elsewhere (e.g., Lahiri (2003), Ghosh (2001)), and sometimes analysed as a synchronic chain shift (Mahanta (2007;2012)). Finally, as mentioned above, Martínez-Gil (2006) discusses the Lena Spanish shift in depth, offering a modified version of Kirchner’s Local Conjunction analysis, which he applies not only to the Lena Spanish shift, but also to a shift in Proto-Spanish (which he is clear applied synchronically, not in distinct diachronic stages) with the form  $\varepsilon \rightarrow e \rightarrow i$ <sup>34</sup>.

Taken together, this is suggestive that there is a tacit acceptance that vowel height shifts are ‘canonical’ examples of synchronic chain shift. This suggests that as a

---

<sup>34</sup> This effect is not listed in the corpus because there are no extant speakers of Proto-Spanish.

starting point for a more detailed study of a potentially unified class of synchronic chain shifts, vowel height is an ideal domain.

The other potential test case would be shifts based on the manner of consonants. Like vowel height shifts, there are well-attested examples of effects of this nature, and they share the necessary characteristic of operating at the level of the segment. Contrasts in consonantal manner have also been used previously in AGL studies (see, for example White (2014)), showing that participants are able to successfully manipulate these kinds of contrast. There are two reasons why I choose to use vowel height as my test case as opposed to these kinds of contrast.

The first is that different consonant shifts in the corpus appear to have different grammatical sources. It is beyond doubt that the putative shifts that are the result of Celtic mutations, for instance, are triggered primarily by morpho-syntactic processes. On the other hand, there are also processes, like that in Gran Canaria Spanish, which do not appear to have any morphological function, and appear to be purely instantiated by their phonological environment, which can be roughly described as V\_\_V. This double dissociation does not exist with regard to vowel height. Whilst it is true that certain vowel height shifts are seemingly triggered by a high vowel (for example, in Lena Spanish or Servigliano Italian), there are other vowel shifts that are not triggered (at least overtly) by the height specification of the following vowel (e.g., Basaa, Gbanu, Guardiaregia Italian). So far, this is very similar to the consonantal manner effects; sometimes a phonological trigger is present and sometimes a trigger is not. The difference lies in the fact that in *all* cases the change in vowel quality is part of the exponence of a particular morpheme. There do not appear to be any purely phonological processes of vowel harmony that work like a chain shift.

The relative coherence of the set of vowel height shifts extends to the nature of the processes that make them up. It is true that some vowel height shifts unambiguously operate purely in terms of vowel height (Lena Spanish being the paradigm case) whilst most others could be argued to be a combination of changes in tenseness and height. However, there are many grammatical approaches (see Kaze (1989), Clements (1991), Parkinson (1996), and, for a more modern take, Trommer (2011))

which envision chain shifts based on vowel height as pure manipulations of one phonological variable, like vowel height, or in Trommer's case, sonority. Even some theorists who use more traditional features such as [+/- ATR] to describe vowel height shifts as a way of representing four-height systems, take the conceptual view that all of the movements in the shift are unified. This feeling is articulated most clearly by Kirchner (1996, p.346), when he sums up the Nzèbi chain shift thus: "Raise a vowel one step from its underlying height value".

It is not even possible to argue that all of the consonant manner effects that are present in the corpus are unified along a similar kind of trajectory. There is little to no unity in the kind of consonant shifts that appear to be able to take the form of chain shifts. In chapter 5 I discuss the two kinds of consonant mutation present in Irish, which have almost entirely different paths (compare the Lenition path,  $p \rightarrow f \rightarrow \emptyset$ , with the Eclipsis mapping,  $p \rightarrow b \rightarrow m$ ). Crosslinguistically, there are other markedly different processes, for example Finnish consonant gradation ( $p: \rightarrow p \rightarrow w$ ). Vowel height shifts show nowhere near this level of variation, and also do not skip steps<sup>35</sup>. As the largest, most coherent group of chain shifts that is present in my survey of such effects, I believe that vowel height alternations are the most sensible choice to serve as a base for my experiments.

### 6.3 Why artificial grammar?

The overarching aim of the following studies is to investigate a) whether chain shifts are learnable at all in some approximation of a synchronic grammar, and b) whether chain shifts are more or less learnable than similar, but minimally different patterns of vowel shift. Question a) has been asked before in relation to specific languages, and indeed even for a vowel raising effect. In a study by Nagle (2013), native Bengali speakers were exposed to a forced choice test in which they were asked to conjugate nonsense verbs. Bengali vowel raising, in the verbal paradigm, is conditioned by certain suffixes, and Nagle's study is, in essence, a wug-test of how productive the vowel alternations in these conjugations are. Nagle found that her subjects applied the correct raising patterns to novel items at a level above chance, but not reliably so, and she acknowledges that her paper is not strong evidence that

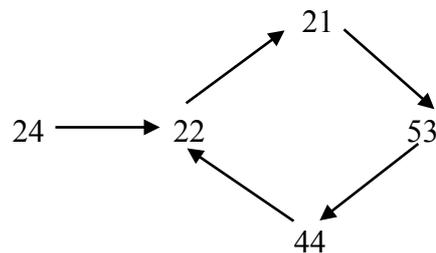
---

<sup>35</sup> Except for the low vowel /a/ in Basaa, which neutralizes with /ɛ/, both being realized on the surface as [e]

the chain shift in Bengali is productive. This is partially because of Nagle’s use of orthographic stimuli, as there are ambiguities in the Bengali spelling system involving vowels that are important in chain shift patterns (Nagle, 2013, p. 20). This merging of symbols for two of the crucial sounds in the chain shift makes it difficult to reliably interpret the answers given that involve these symbols. In turn, it is difficult to make any strong claims about the Bengali shift on the basis of Nagle’s study.

Another wug-test of the productivity of a chain shift, a circular tone sandhi effect in Xiamen, also failed to find evidence for productivity of the shift as a whole (Zhang, Lai & Sailor 2006). In this language, word-final tones appear differently in isolation (in their citation form), and before other words (their sandhi form). These mappings form a circle, as shown in figure 6-1 and the examples in (2) (both reproduced from Barrie (2006, p. 132), who in turn cites Chen (1987)):

Figure 6-1: Xiamen tone circle



(2)	CITATION FORM	SANDHI FORM
(a)	we-24 ‘shoe’	we-22 tua-21 ‘shoe laces’
(b)	wi-22 ‘stomach’	wi-21 pih-22 ‘stomach ailment’
(c)	ts’u-21 ‘house’	ts’u-53 ting-53 ‘rooftop’
(d)	hai-53 ‘ocean’	hai-44 kih-24 ‘ocean front’
(e)	p’ang-44 ‘fragrant’	p’ang-22 tsui-53 ‘fragrant water’

Zhang, Lai, and Sailor’s Taiwanese participants did not show evidence of having internalized the total pattern in any coherent way. However, there was evidence that certain *parts* of the shift were learned significantly more easily than other parts of the shift. This suggests that speakers were capable of extending certain salient generalizations, suggesting in turn that the shift as a whole is not productive.

Quite apart from the issue of whether vowel height and tone circles are in any way related, it should be noted the Xiamen tone circle is the subject of much debate, as certain theorists place circular shifts completely outside the domain of chain shifting, and indeed outside of phonology. Moreton (2004b), for example, asserts that there are no genuine circular chain shifts, and that this is because they are not logically possible within the confines of Optimality Theory. Anything that has the appearance of a circular chain shift, in his view, is either a completely morphological process or has been incorrectly described (though see Barrie (2006), Mortensen (2006) for OT approaches to the Xiamen tone circle). Zhang, Lai, and Sailor (2006, p. 453) also point out that circular shifts provide a problem for traditional rule-based theories, as there is no way to coherently order the necessary rules.

In later work, Nagle (2016) focuses on Zhang, Lai, and Sailor's finding that tone Sandhi was more productively applied when there was some phonetic motivation for that particular part of the Sandhi pattern. Following this thread, her second experiment investigates whether similar co-articulatory effects are present in the vowel raising effect in Bengali. The results of this production experiment suggest that co-articulatory effects are present in Bengali speakers who exhibit the chain shift pattern, but not to a large enough extent to justify a purely co-articulatory explanation for the chain shift. Both of these studies, and wug-tests in general, are limited to testing whether certain patterns that already exist in a language can productively be extended to novel forms. Even if an experiment strongly suggests that a particular example of a potential synchronic chain shift does not appear to be productive, the broadest conclusion that can be drawn is that in that language that particular effect is not productive. It is, therefore, not possible to make a general claim about the reality or otherwise of synchronic chain shifting by wug-testing speakers of a chain shift language.

In order to test a wider typological claim, it is useful to use a paradigm that does not rely so heavily on the language of the participants, or indeed any extant language. For this reason, an Artificial Grammar Learning (AGL) paradigm was used in the following studies. AGL experiments are studies in which participants are introduced to a fragment of an invented language that they have never seen before. Once habituated to the language, participants are generally either tested on whether they

have internalized a particular rule or process that is present in the stimuli they have been exposed to, or to what extent they can generalize a process to unfamiliar items (for a recent review of AGLs in phonology, see Moreton & Pater (2012a; 2012b). Important subsequent studies include Finley & Badecker (2012), Baer-Henney & van de Vijver (2012), White (2013)<sup>36</sup>, and Moreton & Pertsova (2014)).

AGL studies are useful because they allow for complete manipulation of the language fragment that is used. Any variable which one wishes to test can be tested without the influence of the vagaries of natural language. More generally, AGL experiments are a compromise between the trend in linguistics towards detailed empirical studies of natural language, and the more introspective ‘Galilean’ style adopted in classic generative research. Indeed, certain AGL studies (e.g., Becker, Nevins & Levine (2012), Wilson (2003; 2006)) are explicit that they are tests for certain principles of Universal Grammar. Becker, Levine, and Nevins (2010) provides something of a manifesto for this approach, stating that the aim of AGL studies and certain wug-tests is to “make the evidence weigh in favour of UG” (p.2). Whatever position one takes on the specific nature of UG, the advantages of being able to test a theoretical position with real-world speakers who do not exhibit the pattern that is being tested in their native language are clear.

Question b), whether a chain shift effect will be more or less learnable than other, similar patterns, feeds into a major debate in the AGL literature, sparked by Moreton & Pater’s review articles for *Linguistic Compass* (2012a; 2012b). The main claim of these articles is that there are two kinds of cognitive biases that may potentially affect the learnability of a pattern. The first is some notion of phonological complexity. In simple terms, this can be expressed in the amount of features that are manipulated in a particular pattern. Complexity biases are present in many other learning tasks, and Moreton & Pater (2012a) discuss the relationship between phonological complexity and complexity in other domains, finding an instructive analogue in the pattern learning experiments of Shepherd et al. (1961), in which visual pattern recognition was tested. The constituents of Shepard et al’s patterns

---

<sup>36</sup> White (2013) suggests that an AGL based on synchronic chain shifts would be an interesting experiment, though his focus would be on whether participants find it easier to learn chain shifting or neutralizing patterns.

were eight geometric shapes. There were three ways in which the shapes differed from each other (triangles vs. squares, big vs. small, black vs. white). Shepard et al. found that certain simple groupings that relied on a single feature (for instance ‘all of the black shapes’ vs. ‘all of the white shapes’) were easier for participants to learn than groupings that relied on more features (‘the black circles and the white triangles’ vs. ‘the black triangles and the white circles’). This suggests that ease of learning rests on the domain-general finding that simpler patterns are easier to learn than complex ones. Moreton & Pater (2012a) review a large collection of studies showing that this kind of complexity does indeed influence the learnability of particular phonological patterns.

The other kind of bias is based on substance, or naturalness. Certain studies, such as Carpenter (2010), Baer-Henney & van der Vijver (2012), Becker, Nevins, and Levine (2012), and Finley & Badecker (2012) suggest that patterns which are more ‘natural’ than others are easier to learn. This naturalness can take the form of a principle of Universal Grammar, as in Becker, Nevins and Levine (2012), who test whether English speaking participants are biased towards patterns featuring initial syllable protection, a tendency which has been claimed to be universal. Equally, some more detailed phonetic bias may be explored, such as the tendency for front vowels to participate in height harmony triggered by both front and back vowels, as opposed to back vowels, which consistently raise only when the trigger is also back (2012). Finley & Badecker also suggest that “typological tendencies” may play a role in identifying potential biases. Indeed, citing Linebaugh (2007), Finley and Badecker note a typological asymmetry in height harmony systems, in which “at least 52” (Finley & Badecker, 2012, p. 291) languages exhibit the asymmetry discussed above. Only two languages appear to have the reverse pattern, where front vowels do not raise when the harmony trigger is a back vowel, but back vowels are raised whether the trigger is front or back.

The three patterns in the studies below differ in various ways, both in terms of complexity and naturalness. A learnability advantage for one pattern over the other, then, would not just be interesting from the perspective of the synchronic reality or otherwise of chain shifting, but would also add weight to one side or the other of the ‘only complexity’ or the ‘complexity vs. naturalness’ sides of the AGL debate.

## 6.4 The study<sup>37</sup>

This pilot study, then, aimed to show how one could go about addressing two main questions, one general and one specific.

*General question:* Is it possible for participants to learn *any* kind of two-step interaction in an Artificial Learning paradigm?

*Specific question:* Is there any particular learnability advantage or disadvantage to a chain shifting pattern as opposed to minimally different, non chain-shifting patterns?

The more general question is as much methodological as it is theoretical. A potential reason for the lack of productivity in the wug-tests performed by Zhang, Lai, & Sailor (2006) and Nagle (2013) is that chain shifts are complex patterns to learn in the first place. Most artificial grammar learning experiments test the application of either a static pattern (e.g., Carpenter (2010)), or one morphophonemic alternation (e.g., Finley & Badecker (2012), White (2013), though see Ettliger (2008) and Ettliger, Bradlow & Wong (2014)). A key component of synchronic chain shifts is that they constitute (at least) two separate alternations. Therefore, a contribution that these studies make is to develop a rarely used extension of the paradigm, testing whether it is possible to use AGLs to teach multiple, featurally distinct alternations at the same time.

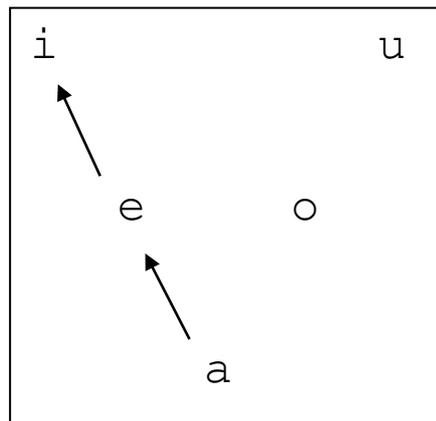
The specific question is slightly more complex. Whatever their source, it is clearly true to say that chain shifts are patterns that reoccur in various forms across the world's languages. It is therefore a pertinent question whether there are learnability advantages to patterns of this nature, as compared to other, equally plausible kinds of effect. To attempt to address this question, participants in this study were taught one of three distinct fragments of what was described to them as an 'alien' language. In the nominal paradigm of this language, there was a three-way split in number between singular, dual and plural forms. Singular stems in the language were always CəCVCə, and the dual and plural forms consisted of a suffix ({-sə} and {-tə} respectively). Crucially, as well as the addition of the suffix, pluralization also conditioned certain vowel changes.

---

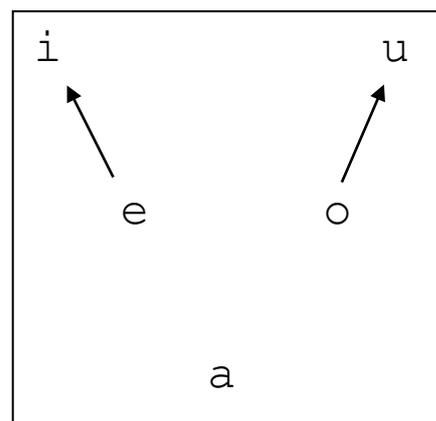
<sup>37</sup> Thanks are due to John Harris, Aditi Lahiri, Andrew Nevins, and Jamie White for many helpful discussions about experimental design. All mistakes, of course, are mine.

Depending on the condition, these vowel changes would result either in a chain shift (low → front mid → front high), a unified raising of both front and back mid vowels to high, or a mixed condition (low → front mid, back mid → back high). Figure 6-2, below, illustrates these patterns:

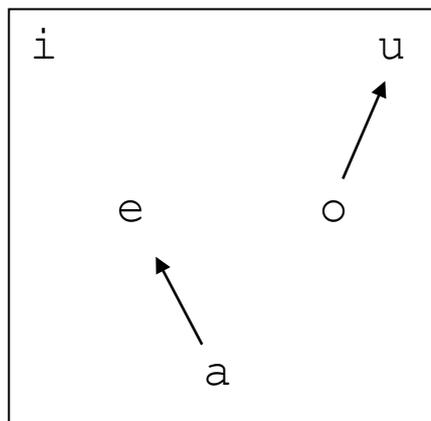
*Figure 6-2: Experiment conditions (arrows show shifts between singular and plural forms. All other vowels do not change between singular and plural).*



*Chain Shift condition*



*Coherent Raise condition*



*Mixed Raise condition*

The study was an online, 2-alternative forced choice study, in which participants would begin by passively learning the pattern through listening to the language, and later progressed to a stage where they were asked to identify which of the words they had heard was an allowable word in the language that they had learned.

#### 6.4.1 Stimuli

80 CəCVCə stem stimuli were created for this experiment<sup>38</sup>, using the phonology of Standard Southern British English as a loose base. It should always be borne in mind that the stimuli created in AGL experiments are not the same as, for example, wug words. In a wug test, it is usually important for control purposes that the stimuli obey all of the inventory and phonotactic constraints of the language, as what is being tested is the productivity of a real pattern of that language. In AGL experiments, deviations from the language of participants are acceptable if they contribute to a clearer overall experimental design. However, it is also important to take into account the phonology of the L1 of the intended group of participants. In order to boost the chances that participants will use their phonology in order to learn the pattern, the language should sound clearly ‘language-like’ and reasonably familiar to participants’ native language(s). That is to say, it should not include sounds that are dissimilar enough from the participants’ native language(s) that they would be encouraged to think of them as non-linguistic.

The stimuli were recorded by a phonetically trained speaker of British English. Recordings were made in a sound attenuated booth using a Rode NT-1A microphone and Adobe Audition 2.0 software. The samples were recorded at 44100 Hz. All of the chosen sound files were normalized for amplitude to an RMS level of 70 dB SPL, but were not altered in any other way. There is no real consensus on the best way of presenting aural stimuli in AGL experiments, with certain researchers using synthesized speech (e.g., Moreton (2008)), some splicing parts together from naturally recorded speech (e.g., Finley & Badecker (2012)) and some using natural speech (e.g., White (2013)). I chose to use natural speech because the results of early pilot experiments using the MBROLA speech synthesizer suggested that participants found it difficult to discriminate between /e/ and /i/ sounds. The consonant inventory of the language was /p t k b d g m n s v r l/. These consonants were chosen to represent a wide range of common English consonants. They were assigned pseudo-randomly to CVCVCV frames, and any duplicate forms were changed by hand.

---

<sup>38</sup> One stimulus, from the set of CəCiCə was excluded during pre-testing, as it was not clearly articulated in the recording.

The alien language has a vowel inventory of six vowels: /i, e, a, o, u, ə/. Whilst it is true that Southern Standard British English does not have monophthongal /o/ or /e/, in a reduced system of this kind, all that was required was that the mid vowels serve two key phonological functions. The first was that it should be clearly mid, as opposed to low or high. The second was that /o/ be clearly distinct from the other vowels in the system, which was tested for in a small AX study (see next section for further details). The five non-schwa vowels occur only in stressed position, and schwa appears only in unstressed position. This asymmetry serves to grant additional salience to the stressed vowel, and should be familiar to English native speakers, as most unstressed vowels in English are realized as schwa.

The overall pattern then, schwa aside, is a standard triangular vowel system, of the kind familiar from languages such as Basque, Czech, Greek, Hebrew, Maori, Tagalog, Tongan, and Spanish. The exact six-vowel system, with schwa, is the system found in, for example, Armenian, Pashto, and Nepali. An early theoretical choice was whether to use this, triangular system, or a quadrilateral system of the kind /i, e, a, ɒ, o, u, ə/. This seven-vowel system had potential methodological advantages over the six-vowel system that was eventually chosen, particularly in terms of counterbalancing, as it was a system with an equal number of low, mid and high vowels. This is particularly salient in the coherent raising condition. Using a six-vowel system, it is only possible to create a language fragment in which mid vowels raise to high. In a seven-vowel system, it would be possible to create a counterbalancing condition in which all low vowels raise to mid. However, there are two problems with a seven-vowel system, one theoretical and one practical.

From a theoretical perspective, it is worth noting that all the languages that feature putative chain shifts, whether they are Bantu, Indo-Aryan, Nilotic or Romance, all have triangular vowel systems. Whilst in most putative chain shift cases in the corpus, the triangular systems that we observe are more elaborate than simple five-vowel systems, a five-vowel system can support a raising chain shift, the classic example being the system found in Lena Spanish. As far as I am aware, there are no synchronic vowel raising or lowering shifts in languages with quadrilateral systems. An experimental paradigm that presented a chain shift pattern within an inventory

that could not potentially house the pattern in natural language is potentially too much of a deviation from phonological reality.

From a practical perspective, it is very difficult to create stimuli for the seven-vowel system described above that are distinguishable enough from one another for English speakers to be able to reliably discriminate all of the relevant contrasts, as shown in the following subsection.

#### 6.4.1.1 An AX test measuring consonant discriminability

28 pairs of stimuli were created from a random subset of the overall set of 96 stimuli that were designed for the seven-vowel version of the experiment. These pairs were minimally different in that the first member of each pair always ended with {-sə} and the second member always ended with {-tə}. C1, C2, and C3 were identical across pairs. Participants were instructed to listen to whether the main vowels (V2) in the pairs of words that they heard was the same or different. Examples of matching and non-matching pairs are shown in table 6-1, below:

Table 6-1: Examples of AX test stimuli

Matching pair	bəsonəsə	bəsnətə
Non-matching pair	ləmpəsəsə	ləmpətə

The stimuli were counterbalanced such that each contrast (for example, /o/ vs. /ɒ/) occurred four times, twice in the order o > ɒ and twice in the order ɒ > o. For each of the six main vowels there were two matching pairs (e.g, o > o). As before, a phonetically trained male speaker of British English recorded the stimuli. I did not discuss the purpose of the experiment with the speaker, but we did discuss the need for the main vowels in the stimuli to be as distinct as possible. The stimuli were concatenated using praat (Boersma & Weenick 2013).

The stimuli were presented online to participants using *Experigen* (Becker & Levine (2010)), a piece of bespoke software designed for phonological experimentation (see, for example, Becker, Nevins, and Levine (2012), Gouskova & Becker (2013), Nagle (2013)). Participants were told to press a play button which triggered one pair of stimuli. One word played directly after the other. Once the stimuli had played,

participants were asked whether the vowels in the two words that they had heard were the same or not. If they believed they were the same participants pressed a button marked ‘Yes’, and if they believed that they were different they pressed a button marked ‘No’. The test was self-paced.

19 participants (9 female, average age 29.6 years) volunteered their time to take part in the study. Native language data was given by all participants (12 British English, 2 American English, 2 Polish, 2 Italian, 1 Swedish, 1 Mandarin). Whilst the target group for the experiment was native English speakers, I did not specify a particular native language for the AX test. In an ideal world, the differences in vowels should have been salient for speakers of all languages. Tables 6-2 and 6-3 show the proportion of *correct* answers given for each of the stimuli pairs, as well as the raw number of errors, across the 19 participants. Figure 6-3 shows the same result, giving the proportions of correct answers by pair:

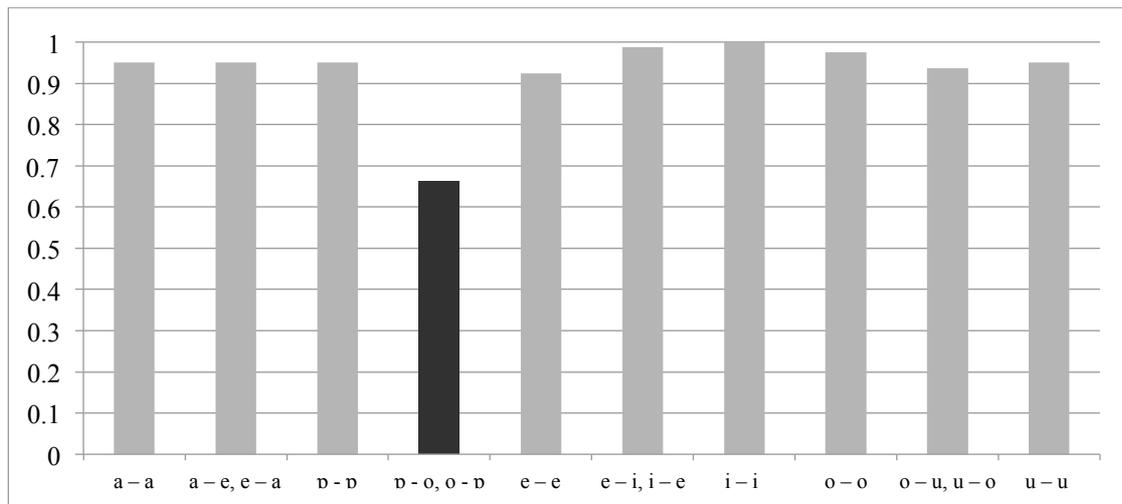
*Table 6-2: Front vowels*

Vowel pair	Proportion Correct (# of correct/total)
a – a	0.95 (38/40)
a – e, e – a	0.95 (76/80)
e – e	0.925 (37/40)
e – i, i – e	0.9875 (79/80)
i – i	1 (40/40)

*Table 6-3: Back vowels*

Vowel pair	Proportion Correct (# of correct/total)
ɒ - ɒ	0.95 (38/40)
ɒ - o, o - ɒ	<b>0.6625 (53/80)</b>
o – o	0.975 (39/40)
o – u, u – o	0.9375 (75/80)
u – u	0.95 (38/40)

Figure 6-3: Proportion of correct answers by vowel pair



As is shown by the bolded, italicized entry in the table above, the only difference which participants were reliably unable to discriminate was the difference between /v/ and /o/. Performance on every other pair was at above 90%. It is plausible to suppose that these pairings are, all else being equal, discriminable in a straightforward way, though the sample is too small for reliable statistical testing.

Easy discriminability is of utmost importance in the AGL paradigm, given that speakers need to not only correctly and quickly identify a set of phonemes that form parts of words that they are not immediately familiar with, but also *manipulate* those phonemes and understand that some of them alter across paradigms whilst others stay the same. For this reason, the six-vowel system was chosen over its rival seven-vowel system.

#### 6.4.1.2 Conditions and predictions

Stimuli were all created in sets of three. These sets of three consisted of a bare stem, a dual form, and a plural form. In the dual form, the suffix {-sə} is added to the stem, and no further alterations are made. In the plural form, the suffix {-tə} is added to every form. At this point, however, additional changes are made depending on the experimental condition.

In the chain shift condition, stems with an underlying /a/ change so that they are realized with [e] in the plural. Simultaneously, stems with underlying /e/ change so

that they are realized with [i] in the plural. Throughout the discussion I will refer to stimuli whose vowels change between singular and plural as ‘dynamic’, regardless of the condition that they appear in. Stems with /i/, /u/ and /o/ do not change in the plural. I will refer to these non-changing stimuli as ‘static’. In the coherent raising condition, stems with underlying /e/ raise and surface as [i] in the plural, and stems with underlying /o/ raise to surface as [u] in the plural. Underlying /a/, /i/ and /u/ surface without change in the plural. Finally, in the mixed raising condition, underlying /a/ raises to [e], whilst simultaneously underlying /o/ raises to [u]. Underlying /e/, /i/, and /u/ surface faithfully. Tables 6-4, 6-5, and 6-6 illustrate the mapping relationships that occur in each condition with examples:

*Table 6-4: Chain Shift examples*

Singular	Dual	Plural
nəmatə [a]	nəmatəsə [a]	nəmetətə [e]
gəsədə [e]	gəsədəsə [e]	gəsədətə [i]
səbilə [i]	səbiləsə [i]	səbilətə [i]
vəgolə [o]	vəgoləsə [o]	vəgolətə [o]
ləkugə [u]	ləkugəsə [u]	ləkugətə [u]

*Table 6-5: Coherent Raise examples*

Singular	Dual	Plural
nəmatə [a]	nəmatəsə [a]	nəmatətə [a]
gəsədə [e]	gəsədəsə [e]	gəsədətə [i]
səbilə [i]	səbiləsə [i]	səbilətə [i]
vəgolə [o]	vəgoləsə [o]	vəgulətə [u]
ləkugə [u]	ləkugəsə [u]	ləkugətə [u]

*Table 6-6: Mixed Raise examples*

Singular	Dual	Plural
nəmatə [a]	nəmatəsə [a]	nəmetətə [e]
gəsədə [e]	gəsədəsə [e]	gəsədətə [e]
səbilə [i]	səbiləsə [i]	səbilətə [i]
vəgolə [o]	vəgoləsə [o]	vəgulətə [u]
ləkugə [u]	ləkugəsə [u]	ləkugətə [u]

The three conditions were designed to differ along three parameters, in order to investigate the relative importance of these parameters. These parameters, and how they are realized in each of the conditions, are shown in table 6-7.

Table 6-7: Comparison of conditions

Condition	Featurally simple/complex?	Coherent/Incoherent?	Transparent/Opaque?
Coherent	Simple	Coherent	Transparent
Chain	Complex	Coherent	Opaque
Mixed	Complex	Incoherent	Transparent

#### 6.4.1.2.1 *Simplicity vs. complexity*

A bias that is widely claimed to be present in language learning is that simple patterns are easier to learn than complex patterns. Moreton & Pater make this point particularly forcefully in their (2012) review of the AGL literature. The idea that features might be a good analogue for simplicity (patterns involving the manipulation of fewer features being simpler) goes back as far as Shepherd et al's (1961) study on visual pattern learning (see section 6.3 for discussion of this point).

In this experiment, vowel height features were the elements of complexity that were being manipulated. The Coherent Raising condition could be simply characterized as 'all mid vowels raise'. This means that, even though in every condition there are two alternations, both of the alternations in the coherent raising condition rely solely on the manipulation of a single feature, [+/-high]. The featural change is exactly the same in the  $e \rightarrow i$  and  $o \rightarrow u$  alternations; [-high]  $\rightarrow$  [+high]. The same analysis is not possible in the chain shift and mixed raising conditions. If either of these patterns are to be learned, participants must manipulate two separate features. In the chain shift condition, the  $a \rightarrow e$  mapping requires a manipulation of [+/-low], whilst the  $e \rightarrow i$  mapping, as in the coherent raising condition, requires a manipulation of [+/-high]. The mixed raise condition is complex in exactly the same way ( $a \rightarrow e$ ,  $o \rightarrow u$ ). The prediction that this make is that, all else being equal, the coherent condition should be easier to learn than the chain shift or mixed raise conditions.

#### 6.4.1.2.2 *Coherence vs. Incoherence: Does the pattern act on a natural class?*

Coherence is a parameter with less of a weight of literature behind it than that of complexity biases, but it is inescapably relevant when considering processes like chain shifts that have been claimed to essentially be functions of their mapping relations. The claim that we should get chain shifts because of certain grammatical principles suggests that these grammatical principles should be able to lend a certain

learnability advantage to patterns of this nature. The notion of coherence that I use is essentially that the segments undergoing the changes can be referred to by reference to one *natural class*. For example, the alternations in the chain shift condition can be referred to as “raising of all vowels with the specification [-back]”. The coherent raise condition can be referred to in equally simple terms; “raising of all vowels with the specification [-high]”. The mixed condition is not amenable to this kind of analysis. A [-back, +low] vowel raises to become [-back, -low], whilst a [+back, -high] vowel raises to become a [+back, +high] vowel. This does not represent an overall effect, rather two separate effects that happen to have similar effects. Therefore, if coherence is an important factor in the relative learnability of the patterns in the experiment, then the prediction would be that the coherent raising and chain shift conditions should be more easily learnable than the mixed raise condition.

#### *6.4.1.2.3 Transparency vs. ‘opacity’*

Another parameter that is manipulated in the experiment is whether the processes that create the overall effect in each condition are transparent or opaque. There has long been an assumption that opaque patterns present a challenge to learners. Baković succinctly sums up this position (whilst, it must be said, critiquing it): “opacity’s original *raison d’être* is Kiparsky’s claim that an opaque rule is difficult to learn” (2011, p. 41). I acknowledge that, in the previous chapter, I asserted that patterns like the Lena Spanish effect (upon which the chain shifting pattern in this experiment is based) may actually be viewed as transparent, and on these accounts cannot involve any derivational opacity. However, this goes directly against how these processes are characterized in most of the literature (see, e.g., Baković (2011), Łubowicz (2003a; 2011; 2012)), where a counterfeeding rule order is presented as the rule-based instantiation of chain shifting. My prediction would lead to a rather negative outcome; in my view all of the interactions can be viewed as transparent, so there would be no learnability advantage across the three conditions. The standard view, by contrast, makes the prediction that the coherent raising pattern and the mixed raising pattern should be more easily learnable than the chain shift pattern, given the derivations that would be required, which I present in table 6-8.

Table 6-8:

*Derivations in the three conditions*

Chain shift condition (opaque, 2 rules)

	/a/	/e/	/i/	/o/	/u/
[-high, -back] → [+high]	-	i	-	-	-
[+low] → [-low]	e	-	-	-	-
	[e]	[i]	[i]	[o]	[u]

Coherent raise pattern (transparent, 1 rule)

	/a/	/e/	/i/	/o/	/u/
[-high, -low] → [+high]	-	i	-	u	-
	[a]	[i]	[i]	[u]	[u]

Mixed raise pattern (transparent, 2 rules)

	/a/	/e/	/i/	/o/	/u/
[+low] → [-low]	e	-	-	-	-
[-high, +back] → [+high]	-	-	-	u	-
SR	[e]	[e]	[i]	[u]	[u]

Overall then, we can present several hypotheses, along with the null hypothesis that there will be no significant differences among the conditions. In all cases, we would expect that the coherent raising condition should be the easiest to learn. It relies on the manipulation of just one feature, affects just one natural class, and is transparent. By all metrics, it is the simplest pattern. It is harder to predict whether it will outperform both the chain shift *and* the mixed raise condition, or just one of the two conditions, though it should always outrank at least one of the others, given that there will always be a metric on which it is simpler. As well as this, basic complexity is assumed to be a much more substantive basis for pattern learning than more subtle naturalness distinctions (see Moreton & Pater (2012a), in particular).

- (3) (a) Null hypothesis: Coherent, Mixed, Chain
- (b) H<sub>1</sub>: Advantage for complexity only: Coherent >> Mixed, Chain
- (c) H<sub>2</sub>: Advantage for coherence: Coherent, Chain >> Mixed
- (d) H<sub>3</sub>: Advantage for transparency: Coherent, Mixed >> Chain

#### 6.4.2 Participants

181 participants were recruited via the web-based crowdsourcing platform *Prolific Academic* (60 for sub-study 1 (19 Chain shift, 21 Coherent Raise, 20 Mixed Raise), 61 for sub-study 2 (21 chain shift, 20 Coherent Raise, 20 Mixed Raise), 60 for sub-study 3 (20 Chain shift, 20 Coherent Raise, 20 Mixed Raise). 102 participants were male, 77 were female, and 2 did not declare this information. Ages ranged from 18-66, with a mean of 33.82 and a median of 32. Prolific allows filtering of participants based on first language, so website users who did not declare that their first language was English were unable to participate in the study. The study included 91 British English participants (I include 1 Scottish English, 1 Scouse English and 1 Welsh English speaker in this), 81 American English participants, 4 Canadian English participants, and 1 New Zealand English speaker. Two of the remaining four participants did not report a native language, though the assumption would be that some variety of English was their native language as otherwise they would not have been able to sign up for the study. The other two report their native language as Lithuanian and Hindi. Again, my assumption is that these speakers are bilingual, as otherwise they would not have been able to sign up for the study. Participants were paid for their time (£2.67 in sub-study 1, £2.00 in sub-study 2, as it was a slightly shorter experiment, and £2.93 in sub-study 3, as it was a slightly longer experiment).

#### 6.4.3 Methodology

Participants followed a link on *Prolific Academic* to an online experiment. This experiment was presented through *Experigen* (Becker & Levine (2010)). Participants read instructions that informed them that they were going to learn a fragment of an alien language. They were told that in the first part of the experiment, they would hear sets of three words, matched to pictures so that they would know what the words meant. These three words would be the singular, dual, and plural forms of the same noun.

The experiment was self-paced, and participants would press a ‘play’ button to trigger the recording of each word. Participants were also asked to repeat each word after it had been played. This was to make the exposure phase, which was essentially passive, a more active process, which would in turn encourage speakers to engage

more with the task (this is a common procedure in AGL experiments, see e.g. White (2014)). The exposure phase consisted of 32 sets of three stimuli (96 tokens total), in two equal blocks separated by a break. Half of these 32 sets of stimuli featured vowel changes between the singular/dual and the plural, and half did not. Table 6-9, below, shows the breakdown of how stimuli were presented in the training phase, by their mapping relation.

*Table 6-9: Number of stimuli presented in the training phase by condition*

Dynamic		Static		
Chain				
8 x a → e	8 x e → i	6 x i → i	5 x o → o	5 x u → u
Coherent				
8 x e → i	8 x o → u	6 x a → a	5 x i → i	5 x u → u
Mixed				
8 x a → e	8 x o → u	6 x i → i	5 x e → e	5 x u → u

Whilst an argument could be made that it would have been useful to have equal numbers of every potential kind of stimulus, what is of paramount importance is that participants were presented with equal numbers of stimuli in which the vowel did change in the plural, and stimuli in which it did not. This is to militate against a bias to either always change or never change when confronted with a choice.

A more serious methodological issue is that the stimuli were completely randomized. This means that there were not discrete sets of ‘training’ and ‘test’ stimuli as is common in experiments of this kind. This meant that not every participant was tested on every item, and that there are not an equal number of observations for each item. The main problem that this causes is that it makes it difficult to run a thorough by-item analysis, which would control for whether there were certain items that biased participants to react a certain way. In any further studies this would have to be controlled for. However, a priori there is no reason to suspect that any particular stem, for which the consonants were chosen completely at random, should necessarily have this bias. In my analysis, I have attempted to construct a model that is reasonably robust in the face of this kind of variation by including random slopes for each item, and specifying the random effects structure as fully as possible. Note that, at a gross level, all participants *are* exposed to the same number of static and

dynamic stimuli. At this larger level, the study is balanced, but at a lower level it is not. This means that there may be certain individual differences that could have been explained by a more controlled study that cannot receive such an explanation here. Furthermore, the total randomization meant that there was not a strict 50/50 distribution of test stimuli where the correct answer was the first putative plural participants heard, or the second. This might mean that a participant who was guessing at a rate of exactly 50/50, pressing button 1 and button 2 an equal number of times, would not actually get a score of 50%<sup>39</sup>.

#### 6.4.3.1 Sub-studies

Participants were randomly assigned to one of three sub-studies, which differed slightly in their training phases. The reason for this was that in pre-analysis of the results of sub-study 1, the results for dynamic mappings in all conditions clustered around chance, suggesting that dynamic mappings in general were too difficult to learn. In order to ameliorate this, I introduced two methodological simplifications to the training phase, discussed below. In the results section, I take the results of these three studies together, as there does not appear to be a significant difference in overall performance from study to study. However, I do acknowledge that this introduces an unwelcome level of noise into the data. A useful step for further research would be to run the study with more participants, as there may be significant differences between the studies that are only visible when the group sizes are larger.

A potential problem with sub-study 1 was the inclusion of the dual stimuli in the training phase. These were included for two reasons. The first was to show more

---

<sup>39</sup> The degree of skewing from chance does not appear to be that large. The table in this footnote shows how far from exact chance (50%) the stimuli that were presented were. A negative number means that pressing the button marked 1 gives a higher chance of getting a correct answer, and a positive number means that pressing the button marked 2 gives a higher chance. Overall, the skewing from chance across the data was less than 1% (0.788):

	Static	Dynamic	Total
Study 1	2.448%	0.388%	1.418%
Study 2	0.080%	-0.427%	-0.173%
Study 3	2.018%	0.220%	1.119%
Altogether	1.515%	0.061%	0.788%

explicitly the directionality of the vowel shifts. Consider a stem with underlying /a/ in the chain shift condition. The singular form would have [a], whilst the plural form would have [e]. A priori, it is equally likely that the morphological alternation is a raising process from /a/ → [e] in the plural, or a lowering process from /e/ → [a] in the singular. This is not crucial, particularly as we have seen that certain theories predict vowel lowering chain shifts (see, for example, Trommer (2011)), but it did make the pattern more explicit. Secondly, the dual forms act as a baffle: something for participants to attend to that would never be tested and was not especially relevant, the dual stems acted to distract participants from the true nature of the task.

However, a potential problem with the use of the dual is that it gives participants more evidence of forms that do not feature vowel shift regardless of the stem vowel. Any generalization about vowel shifting, already hidden to an extent by the half of the plural forms which do not undergo vowel shift, is further hidden by this addition of forms that are very similar (indeed, the suffixes for dual and plural differ by just one feature, [+/- cont]) where there is never any vowel shift. Because the dual and plural forms are not the same morphological paradigm, this fear cannot be couched in terms of a product-oriented generalization, i.e., the concern that participants would be “less concerned with the shape of the base form...and more with creating a product that resembles other words of the same morphological category” (Bybee & Slobin (1982, p. 285); the term ‘product-oriented generalization’ belongs to Zager (1980)). However, one could argue that something similar is happening here. Presented with 96 forms, in which 80 undergo no vowel change, the pattern is not easy to pick out. In the second sub-study, then, there were no duals in the training phase, simply alternations between singular and plural.

Sub-study 3 continued with this methodological simplification by doubling the amount of training given to each participant. The stimuli for sub-study 3 were exactly the same as the stimuli used for sub-studies 1 and 2. Stimuli in the training phase were repeated once each. In this case, however, there were *four* blocks of 16 training stimuli, with the third block a repetition of the first (in a different, random order) and the fourth block a repetition of the second (again in a different, random order). The first two sub-studies had lower amounts of training than is normal in many AGL studies (in White (2014), the exposure phase consists of 72 trials, and in

Finley & Badecker (2012) there are 120). The amount of training was lower in my experiment. This is because studies like those just referenced are poverty-of-the-stimulus experiments where it is imperative that the pattern is learned fully before the test phase begins. However, I felt that it was worth investigating whether varying the training would lead to any great difference in learnability.

After the exposure phase in all sub-studies, participants were informed via written instructions on screen that they would be presented with new words in the language. They would press a play button to hear the first word, and then another play button would trigger *two* words, both potential realizations of the plural. After these words had played, participants were asked which of the two they believed was the correct word in the language that they had learned. They pressed a button marked '1' if it was the first word that they had heard, and a button marked '2' if it was the second word they had heard. All of the words were new in this phase, and again they were presented with 32 sets of stimuli, exactly parallel in composition to those found in the training session. In order to reduce memory load on sub-study 1 participants during this part of the task, duals were no longer included, so participants would see only a singular and a plural form. At the end of the experiment, participants filled out a short demographic questionnaire.

## **6.5 Results and discussion**

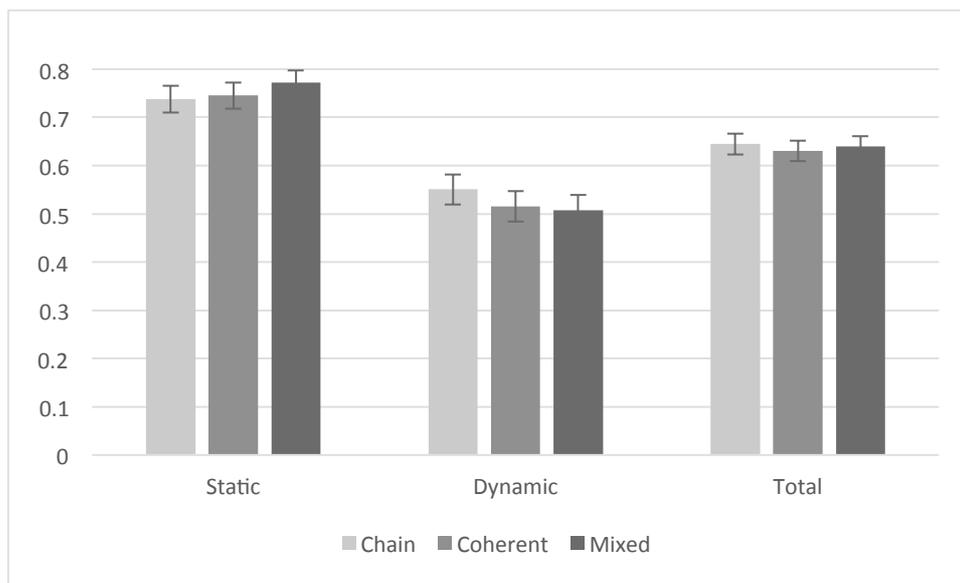
This section begins with the presentation of the total results of the study. This will, in the main, show that the results of the experiment are somewhat flat. There are certain areas of interest that may suggest that there is more occurring than initially meets the eye; I will also discuss this. The overall finding, however, appears to be that there is no great learnability advantage for any of the patterns over any of the others, at least not in a way that is consistent.

Table 6-10 and figure 6-4 show the headline results when all participants are taken together, showing the total scores for static mappings, dynamic mappings, and the learnability of each pattern as a whole.

Table 6-10: Overall results (proportions of **correct** answers)

	Static	Dynamic	Total	<i>n</i>
Chain	0.738	0.551	0.645	60
Coherent	0.746	0.515	0.631	61
Mixed	0.772	0.507	0.640	60

Figure 6-4: Overall proportions of correct results *x* mapping<sup>40</sup>



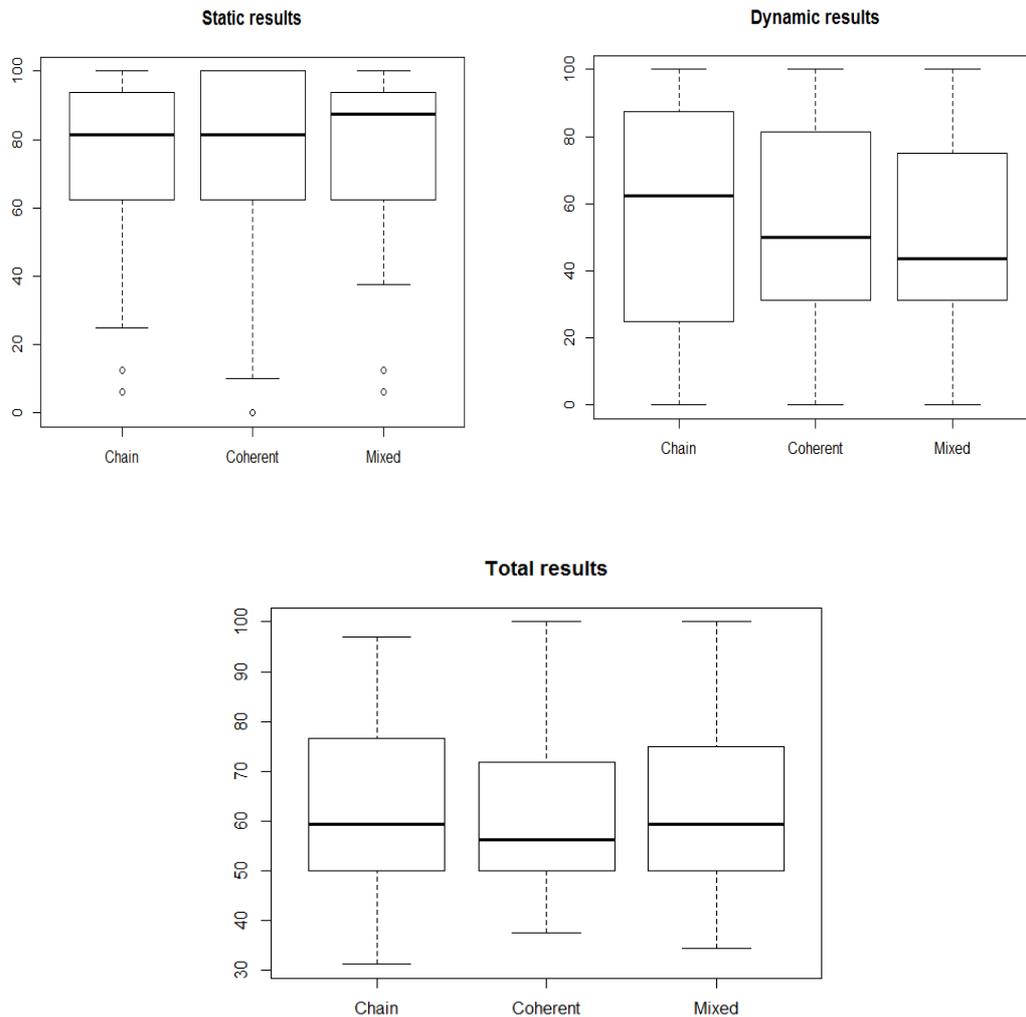
Essentially, there appears to be little difference between the overall scores obtained by participants in any of the three conditions. As previously discussed, it is the *total* score that is of most immediate importance in this study. This is because a participant who scores highly on static mappings but poorly on dynamic mappings has failed to learn the pattern to *exactly the same* extent as a participant who scores highly on dynamic mappings but poorly on static mappings. It is the learnability of the pattern as a whole that is of paramount importance. To this point, between the best learned condition (Chain Shift) and the least well learned condition (Coherent Raising) there is a difference, in terms of correct answers, of just 0.014. There does, however, seem to be a somewhat larger overall difference between the scores for

<sup>40</sup> The error bars in this graph show 95% confidence intervals. I used the Wilson score interval to calculate the confidence intervals, as it recommended by Brown et al (2001) as being a reasonably conservative way of estimating 95% confidence interval scores for proportions. I used an online calculator of the Wilson score interval, it is available at: <http://epitools.ausvet.com.au/content.php?page=CIProportion&SampleSize=320&Positive=146&Conf=0.95&Digits=3>

static mappings vs. the scores for dynamic mappings, across all three conditions. Taken together, the overall proportion of correct results for static mappings (across conditions) is 0.752, and the overall proportion of correct results for dynamic mappings is 0.525.

The bar plot in figure 6-4 appears fairly disappointing for one looking for large differences between the conditions at any level, but it does not give any indication of variance or the spread of results. Even from the bar plot, we can see that the Chain Shift condition had the highest overall score for dynamic mappings and the lowest overall score for static mappings. The reverse is true for the Mixed Raise condition, which had the highest overall score for static mappings and the lowest overall score for dynamic mappings. The potential thus exists for there to have been different overall learning strategies employed by participants in the Chain Shift condition as opposed to the Mixed Raise condition. The boxplots in figure 6-5 illustrate the spread of results more clearly.

Figure 6-5: Boxplots illustrating spread for static, dynamic and total results



The key information that we can take from the boxplots in figure 6-5 is that there is a large degree of variance, in all conditions. It is, of course, problematic that there is something of a ‘floor-to-ceiling’ distribution, particularly in terms of dynamic mappings. Across all conditions, there are participants who were able to essentially answer all questions correctly, and (at least in terms of dynamic mappings) there were participants across all three conditions who answered every question wrong. It is not possible to simply discard these participants, at least not without evidence that they were not completing the task in good faith. The fact that they were unable to learn the pattern is as valid a result as someone who is able to learn the pattern fully, or someone who learns it partially. That said, there are apparent differences. The median score for the Chain Shift condition in terms of dynamic mappings is 0.625,

against a median of 0.5 for the Coherent Raise condition and 0.435 for the Mixed Raise condition. As well as this, whilst the Coherent Raise condition exhibits a clear floor-to-ceiling effect in both static and dynamic mappings, there are far fewer participants operating at floor in terms of static mappings in either the Chain Shift or Mixed Raise conditions.

Before discussing these findings further, it is worth noting a potential problem with the study that had to be tested for post-hoc. As the test was presented online, and a completely smooth playback of the audio files could not be guaranteed, I decided that participants should have the option to repeat sounds that they had not heard the first time. Somewhat naively, I assumed that participants would not use this facility as a chance to repeatedly listen to items as an ad hoc form of additional training. I refer to these participants as ‘multiclickers’ from here on out. *Experigen* records the number of times each sound button has been pressed by each participant, and many participants did in fact press many sound buttons multiple times. This was only discovered in the latter stage of my analysis, and as such it affects all three studies reported here to a greater or lesser extent. To investigate the effects of multiclicking, I set the (admittedly stipulative) threshold that if a participant had multiclicked 10 times or fewer, they would be counted as a non-multiclicker. Overall, 80 participants were multiclickers, leaving 101 non-multiclickers (33 in the Chain Shift condition, 32 in the Coherent Raise condition and 39 in the Mixed Raise condition).

I constructed a binary logistic regression model using the *glmer* function of the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2016) to investigate whether the differences between the conditions were statistically significant. The initial model tested for significant effects of Condition (whether there was a significant learnability advantage or disadvantage for any particular condition), Static vs. Dynamic (whether there was a significant learnability advantage for stems that did not feature vowel shift vs. stems that did, or vice versa), Study (whether the changes made between the three studies significantly affected the overall result), and Multiclicking (whether participants who multiclicked were significantly more successful in the task than those who were not).

Table 6-11: Model featuring full results<sup>41</sup>

Formula: Correct. ~ Condition + Staticdynamic + Multiclick + Study +  
(Staticdynamic + Condition|Stem)+ (Staticdynamic|Participant)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.19967	0.25871	-0.772	0.44
ConditionCoherent	-0.04819	0.2258	-0.213	0.831
ConditionMixed	0.02924	0.2249	0.13	0.897
StaticdynamicStatic	1.47871	0.25345	5.834	5.4e-09***
MulticlickYes	0.26404	0.17831	1.481	0.139
StudyTwo	0.30774	0.21511	1.431	0.153
StudyThree	0.33803	0.21524	1.57	0.116

Certain things are immediately clear from the model. The first is that Staticdynamic is a highly significant predictor in the model, meaning that there seems to be a significant difference between the likelihood of a correct answer if a given stimulus requires a change in vowel quality between singular and plural or not. No other predictor is significant. This is perhaps most important in terms of Condition, which the model suggests is a highly non-significant predictor. To see whether Condition genuinely was a non-significant predictor, I removed it from the model, re-ran the model without it, and then performed an ANOVA comparing the two models, to see whether there was a significant difference in how well they explain the data.

Table 6-12: ANOVA comparing models with and without Condition as a predictor

	Df	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
noconditionmodel	18	5812.8	5932.8	-2888.4	5776.8			
fullmodel	20	5816.7	5950	-2888.4	5776.7	0.1164	2	0.9435

As table 6-12 shows, the model without Condition is almost identical to the full model in terms of goodness of fit. There is no significant difference between the models ( $p = 0.9435$ ). If we compare this model without Condition to a model without Condition and without Study, we can see that Study is also not a significant predictor (Chi sq = 2.9321, Df = 2,  $p = 0.2308$ ). It is worth noting here that a potential problem for my analysis is that, just as in multiclicking, there is a different

<sup>41</sup> I would very much like to thank Steve Politzer-Ahles for his helpful and informative workshop on mixed effects models, and also Bruno Fernandes and Emilia Molimpakis for their help with the statistics. Of course, I take full responsibility for any and all mistakes.

amount of training for participants in sub-study 3 to the participants in sub-study 2 and sub-study 1, so direct comparison between those conditions is potentially helpful. As we can see from the model, the difference between sub-study 1 and sub-study 3 is not significant ( $p = 0.116$ ). It is important to point out that participants did perform better on average in sub-study 3, where they were granted additional training (Estimate = 0.33803, indicating the level to which the log odds of getting a correct answer in sub-study 3 as opposed to sub-study 1 are higher). It is, however, equally germane to note that if the model is re-ordered so that sub-study 2 is the baseline, allowing for direct comparison with sub-study 3, the p-value is so high as to indicate a result that is almost completely identical (Estimate = 0.03028  $p = 0.888$ ).

On this simpler model, Multiclicking also seems to diminish in significance. In the model with just Staticdynamic and Multiclick as predictors, Multiclick is not a significant predictor (Estimate = 0.23252,  $p = 0.191$ ). This seems to suggest that whilst, unsurprisingly, participants who multiclicked are, on average, more likely to give a correct answer to a given stimulus, Multiclicking is not a significant predictor overall.

Evidence from White (2014) suggests these two findings may not be as surprising as they first appear. White's AGL study was a poverty of the stimulus experiment, in which learners were taught a pattern that they then had to productively extend. For this testing to be reliable it was necessary for all participants entering the test phase to have learned the pattern. This took different amounts of time for different speakers, and so different participants received different amounts of training<sup>42</sup>. White built this differential amount of exposure into his model as a predictor, and it was found to be non-significant. This seems to go some way towards explaining why the results of study 3, where the baseline level of training was increased, do not seem to be significantly different to the results of studies 2 and 1. Whilst again acknowledging that comparing the results in this way is a noisy method, it does appear that there is no obvious advantage for participants who have additional training.

---

<sup>42</sup> I acknowledge that this is not a direct parallel, as this differential amount of training was a planned part of White's methodology, rather than a mistake.

Whilst I have addressed the importance of considering the data as a whole, given that there is a significant difference between the static and dynamic mappings, it is worth investigating whether there is an interaction between Static and Dynamic mappings in each of the conditions. I ran a simple model testing for interactions between Condition and Staticdynamic<sup>43</sup>, and none of these interactions were significant ( $p > 0.4$  in all three cases).

The picture when just the overall results are considered seems fairly simple. The only significant predictor of whether an answer to a given question in the test phase will be answered correctly or incorrectly appears to be whether the stimulus requires a vowel shift or not. This accords with the raw results, where in all three conditions there is an advantage for static over dynamic mappings in terms of the proportion of correct responses. What the analysis has not provided so far is a direct comparison of whether Condition had a significant effect within purely static or dynamic mappings. In order to test this, I used very similar models as the model used in the first part of this analysis, but confined the working of the model to either the static or dynamic mappings that were present in the data. The object of this exercise was to more directly investigate whether there were significant differences at this lower level that were washed out in the total analysis.

---

<sup>43</sup> Formula:  $\text{Correct} \sim \text{Condition} * \text{Staticdynamic} + (\text{Staticdynamic} + \text{Condition} | \text{Stem} + (\text{Staticdynamic} | \text{Participant}))$

Table 6-13: Model of dynamic results

Formula: Correct. ~ Condition + Multiclick + Study + (Condition|Stem) + (1|Participant)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.2893	0.3829	0.756	0.45
ConditionCoherent	-0.1172	0.3952	-0.296	0.767
ConditionMixed	-0.2102	0.398	-0.528	0.597
Studytwo	-0.1547	0.3901	-0.397	0.692
Studyone	-0.3445	0.392	-0.879	0.379
Multiclickyes	0.3955	0.3215	1.23	0.219

There are three findings to take from this model. The first is that, at least in terms of the dynamic mapping, there was no reliable effect of multiclicking. That is to say, participants who allowed themselves ad hoc extra training did not perform significantly differently to those who did not (Estimate = 0.3955,  $p=0.219$ ). If Condition is removed from the model, then model fit is essentially identical ( $H_2 = 0.2752$ ,  $Df = 2$ ,  $p = 0.8715$ ). The second finding is that there again appears to be no significant difference between conditions. In fact, the very high p-value suggests that the overall scores between conditions are nearly identical. The third finding is that the two attempts to simplify the study, first by taking out the dual stimuli (sub-study 2) and then by extending the training period (sub-study 3) had no significant effect. That the increase in training was ineffectual is perhaps unsurprising given that multiclicking was not a significant predictor as these participants had given themselves extra training but did not perform significantly better than their fellow participants who did not. Turning again to the static conditions, we see very similar results. The model dealing only with static stimuli is not significantly different if Multiclicking ( $H_2 = 0.3834$ ,  $Df = 1$ ,  $p = 0.5358$ ), Condition ( $H_2 = 0.6596$ ,  $Df = 2$ ,  $p = 0.7191$ ), or Study ( $H_2 = 1.8934$ ,  $Df = 2$ ,  $p = 0.388$ ) are removed from it.

### 6.5.1 Results for non-multiclickers

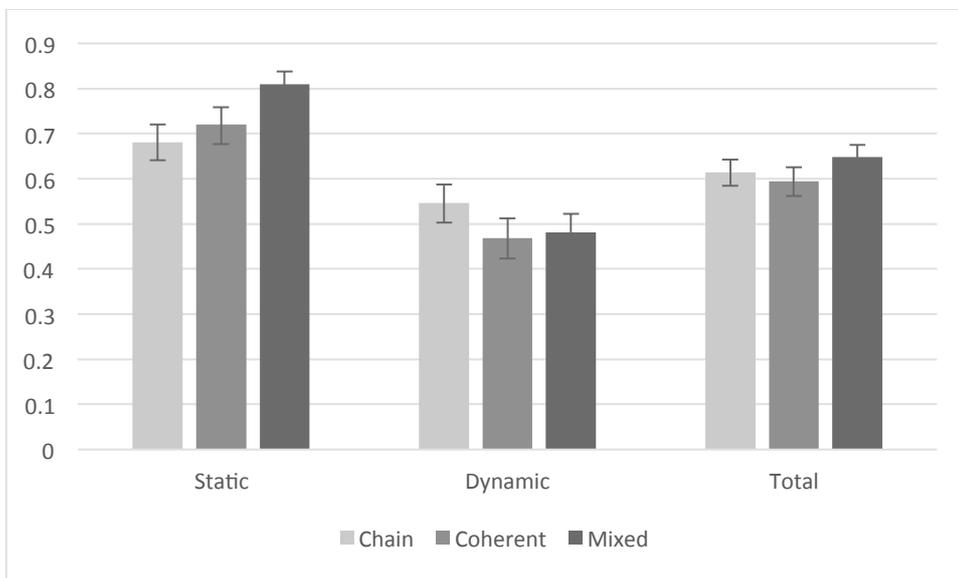
Given that multiclicking is an unintended consequence of the experiment, as opposed to a manipulation that I purposefully added, I am not unsympathetic to the argument that multiclicking participants should be removed as a matter of course, even though there does not appear to be a significant difference in scores between multiclickers and non-multiclickers. It can certainly be argued that the non-multiclick results

present a more homogenous group in terms of how the task was performed. For that reason, I give a full analysis of the data for just the non-multiclick participants here. Table 6-14 and figure 6-6 show the raw results for the non-multiclickers.

*Table 6-14: Average scores for non-multiclickers*

	Static	Dynamic	Total	n
Chain	0.681	0.546	0.614	33
Coherent	0.72	0.468	0.594	29
Mixed	0.809	0.482	0.649	39

*Figure 6-6: Non-multiclick results*



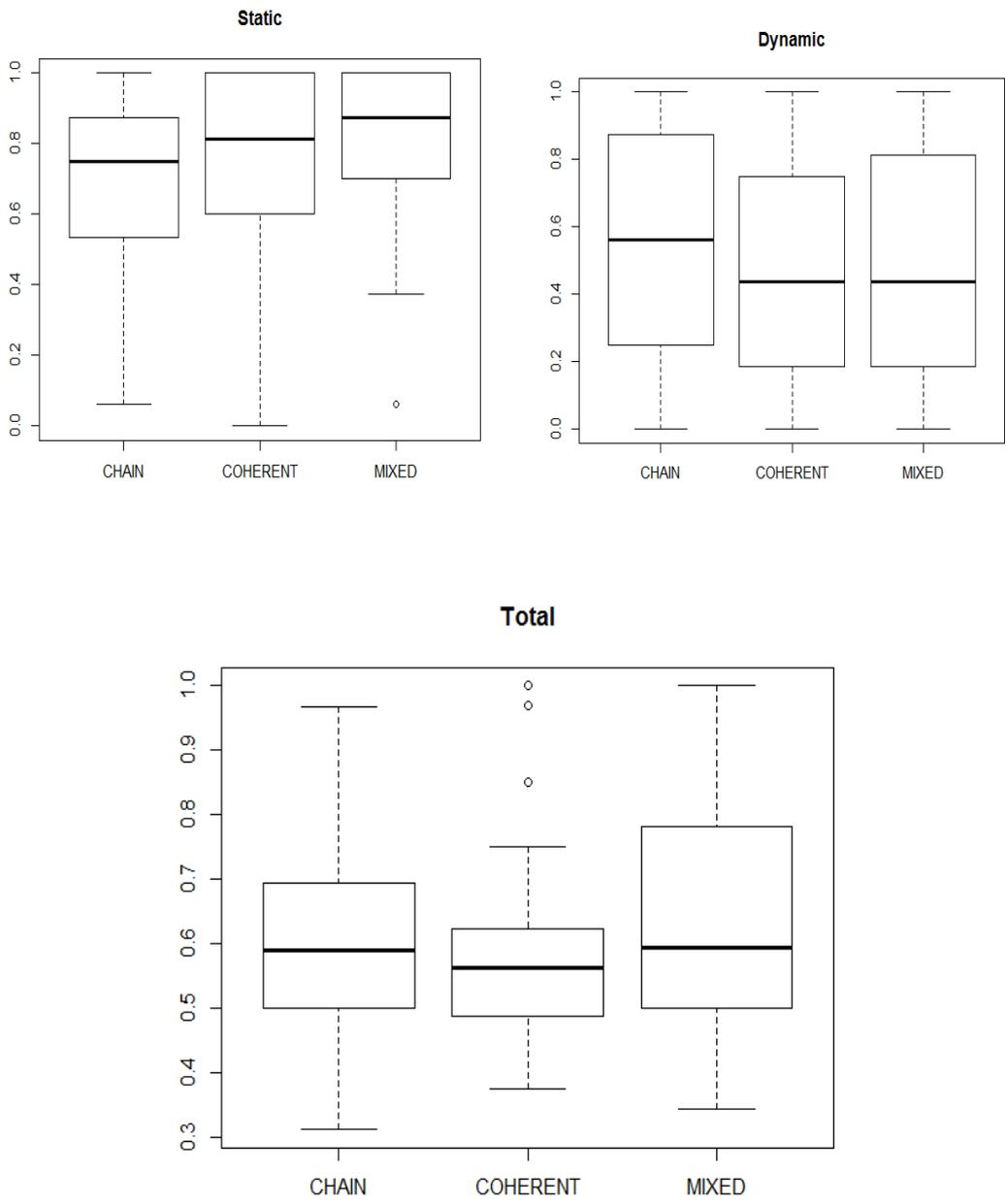
In terms of the ordering relations between conditions, there are two main changes. The first is that, in terms of dynamic mappings, the score for participants in the Coherent Raise condition has fallen slightly below the scores for those in the Mixed Raise condition. This is behind the change in ordering that we observe in the overall totals. The Coherent Raise condition is still the lowest overall score, but the chain shift and mixed conditions have switched over. It should still be borne in mind that the margins of difference are still very small. Table 6-15 shows the difference between the multiclick and non-multiclick participants by condition. Differences preceded by a minus are those in which those participants who multiclicked outperformed their more well-behaved fellows. Differences preceded by a plus are those in which the reverse is true, and non-multiclickers outperformed multiclickers:

Table 6-15: Differences in overall score between multiclickers and non-multiclickers

	Static	Dynamic	Total
Chain	-0.057	-0.005	-0.031
Coherent	-0.026	-0.047	-0.037
Mixed	+0.037	-0.025	+0.009

It is hardly a surprising finding that multiclickers, by and large, outperform their non-multiclicking co-participants in most cases, as multiclickers essentially had access to additional training. It is worth noting that, again, the differences between multiclickers and non-multiclickers are not large (the largest is the 5.7% difference for static mappings in the chain shift condition). Still, this is a purely descriptive generalization. It is vital to investigate whether non-multiclickers behaved significantly differently to their multiclicking counterparts. A first way of looking at this is presented in the boxplots in figure 6-7.

Figure 6-7: Boxplots illustrating Static, Dynamic and Total scores for non-multiclickers



Once again, when we consider the dynamic mappings, there is a clear floor-to-ceiling pattern in each of the three conditions. It appears that both multiclickers and non-multiclickers varied widely in performance when it came to items where vowel shifts occurred, regardless of the nature of the shifts in question. In terms of the static mappings, we see a clear ceiling effect in the Coherent and Mixed Raise conditions, but a wider range of responses in the Chain Shift condition, a point that will be taken

up again later. The main difference when we separate out the multiclickers appears to be that there are fewer participants in the coherent condition who are operating at ceiling, compared to both Chain Shift and Mixed Raise participants in the non-multiclickers. An overall point that I make at this juncture is that an expectation of *all* of the alternative hypotheses for the experiment was that the Coherent Raise condition should be the easiest condition to learn. This has not been borne out at all in the results. In terms of overall score, whether or not mutliclicking is taken into account, participants in the Coherent Raise condition are the *worst* performing participants. It should be noted that whilst one can attempt to explain why this might be, the differences between the three conditions are so small that there may well not be a principled reason for this ordering. In other words, it may just be random variation. Statistical analysis of the data for the non-multiclickers as a whole is presented in table 6-16:

*Table 6-16: Overall statistical analysis of non-multiclicking participants:*  
 Formula: (Correct. ~ Staticdynamic + Study + Condition + (Staticdynamic + Condition|Stem) + (Staticdynamic|Participant))

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.158158	0.33301	-0.475	0.635
Staticdynamicstatic	1.632476	0.359037	4.547	5.45e-06***
Studyone	-0.132838	0.29185	-0.455	0.649
Studytwo	0.174618	0.283147	0.617	0.537
ConditionMixed	0.316893	0.305692	1.037	0.3
ConditionChain	-0.008084	0.306233	-0.026	0.979

As with the overall results, the only significant predictor in the model is Staticdynamic. This means that, overall, a correct answer is significantly more likely if the stimulus in question is a static mapping. Both Study and Condition are highly non-significant predictors. If either one is removed from the model, then there is no significant improvement in model fit (Condition:  $H_2 = 1.4883$ ,  $Df = 2$ , Study:  $H_2 = 0.4751$ , Study:  $H_2 = 1.0619$ ,  $Df = 2$ ,  $p = 0.588$ ). Because of the strongly significant difference between static and dynamic mappings, there is justification for investigating whether this difference persists across conditions. Therefore, a model was constructed (without Study, as there is no reason to suppose that it is a significant factor), examining the interactions between Staticdynamic and Condition:

Table 6-17: Modelling the interaction between static and dynamic mappings amongst non-multiclickers

Formula: Correct. ~ Condition\*Staticdynamic + (Staticdynamic + Condition|Stem) + (Staticdynamic|Participant)

	Estimate	Std. Error	z value	Pr(< z )
(Intercept)	-0.1911379	0.3678374	-0.52	0.6033
ConditionChain	0.5154557	0.5344453	0.964	0.3348
ConditionCoherent	0.0002978	0.55644288	0.001	0.9996
Staticdynamicstatic	2.2909205	0.5663637	4.044	5.2552-05***
ConditionChain:Staticdynamicstatic	-1.5016934	0.8200083	-1.831	0.0671
ConditionCoherent:Staticdynamicstatic	-0.5644481	0.852843	-0.662	0.5081

The chain shift condition exhibits the strongest interaction effect out of the three conditions. The interaction does not quite reach significance but the fact that it is stronger than in the other two conditions accords with the finding that the difference between static and dynamic scores in the Chain Shifting condition (15.5%) is somewhat smaller than this difference in the Coherent Raise condition (25.2%) and smaller still than the difference in the Mixed Raise condition (32.7%). This is an intriguing finding, as it suggests that there is some substantive difference in how participants are responding to the stimuli across conditions. This suggests that it is important again to examine the static and dynamic results, in order to assess whether there is a significant difference between the results at this lower level.

A model testing purely for an effect of Condition in terms of dynamic results performs no better than a model that does not include the effect of Condition ( $H_2 = 0.9603$ ,  $Df = 2$ ,  $p = 0.6187$ ). This suggests that there are no significant differences between the overall scores for Condition. Given that, as in the results where multiclickers were not discarded, we observe a floor-to-ceiling distribution across all three conditions, this is unsurprising. With the large amount of variation that we see in the data, it is not possible to claim with any conviction that the differences we see in the dynamic results (with the Chain Shift scores slightly higher than those in the Coherent Raise and Mixed Raise conditions) are due to anything other than chance, and the statistics give us no reason to suppose that this is not the case.

When we consider the differences in scores in terms of static mappings, however, we see a slightly different picture emerge. Table 6-18 shows the model of the static results:

*Table 6-18: Static mapping results for non-multiclickers*

Formula: Correct.  $\sim$  Condition + Study + (Condition | Stem) + (1 | Participant)

	Estimate	Std. Error	z value	Pr(< z )
(Intercept)	0.7575	0.4689	1.616	0.106
ConditionCoherent	0.4488	0.5086	0.882	0.378
ConditionMixed	1.0479	0.4858	2.157	0.031*
Studytwo	0.5965	0.494	1.208	0.227
Studythree	0.4624	0.4836	0.956	0.339

There is a significant difference between the static scores in the Mixed Raise condition and the Chain Shift condition (which is the baseline in the model). The comparison that we do not get from this model, whether there is a difference between the Mixed Raise and Coherent Raise conditions, is not significant (Estimate = -0.5991,  $p = 0.234$ ). This means that the only significant difference is between the Chain Shift and the Mixed Raise conditions.

It is first important to put this difference into perspective, given the dearth of significant results that the study has so far uncovered. On this score it is important to bear several things in mind:

- 1) The level of significance is not that great ( $p = 0.031$ ). In a more well-controlled study this would not be of such concern, but given that the model is based on a reduced population (because of multiclicking) spread out over three sub-studies and with unequal numbers of items (because of the full randomization discussed in the methodology section), significant differences at this level should be treated with a degree of caution. Additionally, the random effect structure of the model is fairly bare. Barr et al (2013) point out that if random effect structures are insufficiently fully specified, then the risk of Type I error increases to a fairly high degree. However, more complicated random effect structures either had little to no effect on the model or caused the model to fail to converge.

- 2) The static mappings are less important than the overall mappings. The overall mappings show how well the entirety of the pattern is learned. As I have previously pointed out, someone who scores 50% overall can be said to have learned 50% of the pattern. That is equally true whether they have learned all of the static mappings and none of the dynamic mappings, all of the dynamic mappings and none of the static mappings, or some combination thereof.

These caveats aside, it is still important to investigate potential reasons for any significant differences we observe in the data, particularly given that there is no a priori reason why any of the three sets of static mappings should be any harder to learn than any of the other three. The brief AX test described in section 6.4.1.2 indicated that all of the contrasts between vowels in the six vowel system were reasonably salient, hinting that there were no serious problems in the perception of any of the vowels that would suggest confusion on this score.

A similarly small-scale control study was conducted to test this more directly with the full set of test items, to assess whether there was a temptation on the part of participants to shift particular vowels even when there was no training telling them that they should do so. Ten English native speakers were recruited for this study via *Prolific Academic*, though one set of results had to be discarded as the participant had previously participated in one of the other sub-studies. This left nine participants, who were paid £2.00 for their time. The task in the control experiment was exactly the same experimental task as that in the main study, but in this case all of the vowels had static mappings between singular and plural. That is to say, no vowel shift occurred in the experiment, which existed purely to investigate whether participants could reliably a) hear the contrasts between the vowels clearly and b) understand the basic mechanics of the task. The results of this task are displayed by vowel in table 6-19.

Table 6-19: Overall results from small control study by vowel

Vowel	Proportion correct
/a/	0.981
/e/	0.944
/i/	0.907
/o/	0.852
/u/	0.963

This control study was largely performed as a sanity check, and is not large enough to bear statistical comparison with the data from the main study. It would be helpful in future to run a larger control study to ensure that all of the contrasts between vowels are perceived with the same relative ease, but the initial pilot study gives no real cause for concern on this score. The score for /o/ is slightly lower than the other vowels, but it would perhaps be dangerous to read too much into this, particularly when we consider that the difference in actual correct answers between /o/ and the next lowest scoring vowel /i/ is 3 (46/54 vs. 49/54). Whilst without a larger study we cannot completely put to bed the idea that there may be some vowels that participants choose to change more than others irrespective of whether the task mandates that they should, the brief control study here does not point to strong evidence that this is the case.

Why, then, do we observe this asymmetry between the Mixed Raise condition and the Chain Shift condition in terms of static scores? The most interesting potential hypothesis for the purposes of this experiment is that it shows that different strategies are consistently being employed by participants in these two conditions. The data shows that, whether or not multiclickers are included, the Chain Shift condition shows the highest scores for dynamic mappings and the lowest for static mappings. By contrast, the Mixed Raise condition shows the highest scores for static mappings and the lowest score for dynamic mappings. A participant who scores highly on static mappings and poorly on dynamic mappings can be said to be utilizing a conservative strategy. That is to say, these participants do not change *any* vowel, regardless of whether the training phase has taught them that they should. A pattern that is difficult to learn might result in more participants adopting a conservative strategy. We can only adduce descriptive, indirect evidence for whether this is likely to be the case, but looking at how many participants adopted a

conservative strategy in each condition can give us an indication of whether this is a worthwhile direction to pursue. Table 6-20, below, shows how many participants in each condition (of the non-multiclickers) adopted a conservative strategy, here defined as scoring above 80% correct for static mappings and below 20% for dynamic mappings.

*Table 6-20: Conservative participants by Condition*

Condition	Conservative participants	% of participants
Chain	5/33	15.15
Coherent	9/29	31.03
Mixed	9/39	23.08

Proportionally, there is a fairly small difference between participants in the Chain Shift and Mixed Raise conditions (7.93%). This difference is made smaller still if we consider that there was another, equally simple strategy that participants could have employed. This can be referred to as anti-conservativity, and is simply the reverse of a conservative strategy, where a participant changes *every* vowel regardless of whether they have been trained on it being static or dynamic. There were three participants who took this approach in the Chain Shift condition, two in the Coherent Raise condition, and one in the Mixed Raise condition. If we consider all strategic participants together, then we get the results that I present in table 6-21.

*Table 6-21: Conservative and anti-conservative participants by condition*

Condition	Strategic participants	% of participants
Chain	8/33	24.24
Coherent	11/29	37.93
Mixed	10/39	25.64

If we consider strategic participants as being any participant who was at ceiling for one set of mappings and floor for another, then, we can see that participants in the Chain Shift and Mixed Raise conditions are almost exactly as likely as one another to adopt such a strategy. We are then left with a puzzle of why so many more participants in the Coherent Raise condition (proportionally) appear to be strategic in this way. However, it would be unwise to speculate too much on this. The reason for investigating these kinds of strategies at a descriptive level was because of a genuinely significant difference between the scores in the Mixed Raise and Chain

Shift conditions. Given that there is no such significant difference in static scores between the Coherent Raise condition and either of the other two conditions, there is no real justification for assuming that the higher level of participants utilizing a strategy in the Coherent Raise condition is anything other than chance. Indirect evidence for the fact that strategies hold fairly constant across conditions can be found if we consider the data overall, adding the multiclickers back into the analysis. Table 6-22 shows the overall results if we make this move:

*Table 6-22: Strategic participants from complete results (multiclickers and non-multiclickers)*

	Conservative	Anti-Conservative	Total
Chain Shift	11/60 (18.34%)	4/60 (6.67%)	15 (25%)
Coherent Raise	13/61 (21.31%)	2/60 (3.28%)	15 (24.59%)
Mixed Raise	12/61 (19.67%)	2/60 (3.34%)	14 (23.34%)

If we consider the total results, two things become apparent. The first is that, in the Chain Shift and Mixed Raise conditions, the proportion of strategic participants holds almost entirely constant when just the non-multiclickers or the entire dataset is considered. The second is that, when all of the data is considered, the potentially interesting difference between the Coherent Raising condition and the other two conditions is washed out entirely. This lends further grist to the mill of the idea that the difference is motivated by nothing other than chance.

A more likely potential reason for the significant difference in static scores for the Mixed Raise and Chain Shift conditions lies in the treatment of individual vowels. Table 6-23 illustrates the score for each individual vowel across each of the three conditions (amongst the non-multiclickers):

*Table 6-23: Overall scores per vowel for non-multiclickers*

Vowel	Chain Shift		Coherent Raise		Mixed Raise	
	Correct	Static/Dynamic?	Correct	Static/Dynamic?	Correct	Static/Dynamic?
a	0.563	Dynamic	0.678	Static	0.468	Dynamic
e	0.53	Dynamic	0.478	Dynamic	0.764	Static
i	0.782	Static	0.766	Static	0.808	Static
o	0.661	Static	0.457	Dynamic	0.497	Dynamic
u	<b>0.582</b>	Static	0.724	Static	0.856	Static

Perhaps the most striking finding from the table is the very low score for static /u/ → [u] mappings in the Chain Shift condition (italicised in the table). The score for /u/ is almost identical to the scores for the dynamic mappings /a/ → [e] and /e/ → [i]. The difference between the highest dynamic mapping (/a/ → [e]) and the lowest static mapping is just 1.9%. The next smallest difference is *also* in the Chain Shift condition, and is between the (/a/ → [e]) mapping and the (/o/ → [o]) mapping. This difference is 9.8%. Statistics on whether these individual differences are significant would draw on samples too small and unbalanced for the data to be trustworthy. However, statistics on the larger question of whether there is a significant difference between static and dynamic mapping in each condition, illustrate that Staticdynamic is a significant predictor in the Coherent Raise (Estimate = 1.8253, p = 0.0127) and Mixed Raise conditions (Estimate = 2.1668, p = 4.99e-05), but not in the Chain Shift condition (Estimate = 0.8185, p = 0.179).

The question, then, becomes why participants in the Chain Shift condition performed so poorly on the static stimuli relative to the other two groups. Despite the fact that there is a numerical advantage for dynamic mappings in the Chain Shift condition over the other two conditions, we cannot fairly say that this is related to the poor performance on static stimuli, as the difference in dynamic mappings across conditions is not even close to significance. Therefore, we must consider alternative explanations. It is perhaps worth pointing out that there is one static mapping that Chain Shift participants perform roughly as well on as participants in the Coherent Raise and Mixed Raise conditions; /i/ → [i]. The problematic static mappings for those in the Chain Shift condition are those mappings which involve back vowels. As I have pointed out, there does not seem to be a general aversion to back vowels. Indeed, in the Mixed Raise condition, /u/ → [u] is the mapping on which participants score most highly (85.6% correct responses). This further speaks against the idea that /u/ → [u] is an intrinsically difficult mapping, or that there is some intrinsic difficulty with back vowels as opposed to front vowels.

A potential reason for low scores on the low back vowels is related to the nature of the task. The Chain Shift condition is the only condition in which there is only one termination point. In the Coherent Raise condition, both /e/ and /i/ in the singular

map to [i] in the plural, and both /o/ and /u/ map to [u]. In the Mixed Raise condition, both /a/ and /e/ map to [e], whilst again both /o/ and /u/ map to [u]. This is not the case in the Chain Shift condition, in which /a/ maps to [e] and both /e/ and /i/ map to [i]. During the training phase, participants will hear more tokens of sounds that are termination points of shifts. As well as this, it is worth noting that, in all conditions, when participants were presented with static /u/ stems in the test phase, the choice was always between a plural featuring [u] and a plural featuring [i]. In the Chain Shift condition, then, it is conceivable that the increased frequency of [i] in the training phase may have led participants to wrongly learn that [i] was the correct answer more often than it actually was.

If this were the reason for the decreased performance in /u/ stimuli in the Chain Shift condition, then the obvious conclusion to draw would be that this constitutes a task effect, rather than reflecting some phonologically motivated distinction. As discussed in the methodology section, the decision was made to have exactly equal numbers of static and dynamic stimuli in the training phase so that participants would not be biased to change or not change vowels. Whilst I still believe that this was a sensible decision, it leads to a slight unbalancing in terms of how often participants hear certain sounds that might have caused this difference between static scores in the Chain Shift and Mixed Raise conditions. I note that even in the seven-vowel system that was considered in the design phase of the experiment, this problem would still be present. It is an inalienable property of the Chain Shift condition that it will have fewer termination points across a system than the Coherent Raise or Mixed Raise conditions.

A final experimental issue to briefly discuss is that the advantage for static mappings in the Mixed Raise condition seems to come almost completely from sub-study 1. The results for this sub-study (including both multiclickers and non-multiclickers) are shown in table 6-24.

Table 6-24: Sub-study 1 results

	Static total	Dynamic total	Total	<i>n</i>
Chain	0.670	0.533	0.601	19
Coherent	0.655	0.509	0.582	21
Mixed	0.819	0.463	0.641	20

In the subsequent sub-studies, this result is not replicated (in sub-study 2, participants in the Coherent Condition achieve the highest scores for static mappings, and in sub-study 3, there is very little difference between any of the three conditions). The difference between scores for the Chain Shift and the Mixed Raise conditions approaches significance here (Estimate = 1.135,  $p=0.0502$ ). However, even in this sub-study it is worth noting that the amount of participants essaying either conservative or anti-conservative strategies gives no support to the notion that there is a reliably different learning strategy being employed by participants in the Mixed Raise condition. There are five conservative/anti-conservative participants in the Chain Shift condition (four conservative, one anti-conservative), five in the Coherent Raise condition (three conservative, two anti-conservative), and four in the Mixed Raise condition (all conservative).

Additionally, the modifications made to the tasks in sub-studies 2 and 3 were made to all conditions, and should have affected all conditions equally. I cannot think of a reason why the addition of the dual in sub-study 1 would have benefitted participants in the Mixed Raise condition and not participants in the other two conditions. A potentially useful follow-up study would be to simply run more participants in sub-study 1, in order to see if the slightly different methodology does reliably tease out this difference. However, my intuition, given that the methodological alterations made in sub-studies 2 and 3 seemed to make very little difference to the overall scores, is that the most plausible reason for the particularly good performance of Mixed Raise participants (or the particularly poor performance of participants in the other conditions) is a matter of chance. Also, given that, even in sub-study 1, the scores for dynamic mappings and the overall mappings are very similar suggests that any overall result that is lurking in the patterns will require a more substantial overhaul to the methodology of the experiment than simply running more participants (some ideas in this vein are discussed in section 6.7).

### 6.6 More general potential reasons for the null result

However one looks at the data, then, it seems that there is no significant learnability advantage for any particular condition over any other. There are several general methodological reasons why this might be so. As mentioned in the experimental design section, this experiment differs from the majority of AGL experiments in that (in at least some conditions), there are multiple mappings that need to be learned that manipulate different features. In most AGL studies, there is one particular kind of alternation that is learned. In poverty-of-the-stimulus studies, like White (2014)<sup>44</sup>, new stimuli are introduced which are *potential* undergoers of the pattern that has been learned.

In Study 1 of White's experiment, participants were trained on patterns in which either voiceless plosives surfaced as voiced fricatives (which White terms the 'potentially saltatory condition') or voiced plosives surfaced as voiced fricatives (the 'control' condition). Participants were then introduced to segment-types they had not been trained on (voiced plosives and voiceless fricatives in the potentially saltatory condition, voiceless plosives and voiceless fricatives in the control condition). These new stimuli are the crucial data, as White was interested in how and whether the patterns taught in the training phase would be extended. For our purposes, the most important feature is that the two patterns that are taught are one-step patterns. It is true that these patterns differ in complexity: in the potentially saltatory condition where, for example,  $p \rightarrow v$ , there are two features being manipulated ([+/- cont], [+/- voice]); in the control condition, where  $b \rightarrow v$ , one feature is being manipulated ([+/- voice]).

Participants in White's study only ever need learn one rule, even though that rule might differ in complexity. In the Chain Shift and (especially) the Mixed Raise conditions, two rules must be learned that are independent of one another. In the chain shift condition, participants must learn that low vowels become front mid, whilst front mid vowels become high, and in the mixed raise condition, low vowels also become front mid whilst back mid vowels become high. These are not the kinds

---

<sup>44</sup> I use White's study as an example because of its currency and admirable clarity. The classic of this kind of AGL study is Wilson (2006), and there are a great many others listed in Moreton & Pater (2012a; 2012b).

of alternations that logically follow from one another, so the whole pattern must be taught during the training phase. It may be the case that learning two separate alternations of this kind, particularly in the fairly short space of time afforded by the training phase (32 stimuli in sub-studies 1 and 2, 64 stimuli in sub-study 3), represents too difficult a task for the majority of participants.

This may seem an odd place to argue from, given that in all three conditions there were participants who performed completely at ceiling. For these participants, then, the task was clearly not too difficult; if anything, the reverse was true. At this point, it is crucial to mention that there do exist studies in which participants are able to learn complex patterns. Ettliger (2008) and Ettliger, Bradlow & Wong (2014) are studies in which two distinct patterns, one simple and one complex, are taught to the same participants during the training phase. Simple is taken to mean the kind of alternation that has a clear, consistent phonological trigger. The example the authors give is voicing alternations in the English plural. Complex, by contrast, is what at least appears to be a contextless process, where the underlying form of some sound surfaces as something else. In the study above, then, both alternations are what Ettliger, Bradlow and Wong would term Complex. In both Ettliger (2008) and Ettliger, Bradlow and Wong (2014), a subgroup of participants (who Ettliger, Bradlow, and Wong term 'learners') are able to learn both the simple and complex patterns, suggesting that, at least for some participants, learning two rules in an AGL context, and learning complex rules, is possible. This finding is borne out in the current study, as a similar subgroup exists in my experiment. The question, though, is what are these participants learning, and how are they learning it?

In Bradlow, Ettliger and Wong's study, participants were tested not just on the artificial language designed by the authors, but also a battery of standardized memory tests, testing procedural memory, declarative memory and working memory. They found that participants who were able to learn both the simple and complex pattern in the artificial language they had created scored significantly higher on tests of declarative memory than participants who learned only the simple pattern, but that there was no such difference for procedural memory or working memory. Procedural memory is often thought to be closely involved with grammatical rules, a position argued for by Ullman et al in their statement

“grammatical rules are processed by a frontal/basal ganglia “procedural” system” (1997, p. 266). Declarative memory is thought to be used extensively in the early stages of L2, where new words (and even phrases) are often learned *en bloc* (see, e.g., Ullman (2005)). Given that the non-words in the test phase of the current study, and indeed in Bradlow, Ettliger, and Wong’s study, were not words participants had heard before, it is important to consider what other role declarative memory could be playing. Bradlow, Ettliger and Wong speculate that “an alternate hypothesis is that declarative memory may be storing morphophonological patterns” (2014, p.822-823). This suggests that complex morphophonological rules of this nature were learned in a different way to simple phonologically motivated rules, which could be learned by participants who scored highly on tests of procedural memory but comparatively poorly on tests of declarative memory. Bradlow, Ettliger, and Wong suggest that the learning that is taking place is potentially learning by analogy, rather than rule-based learning.

Whilst I did not test the relative memory skills of my participants, I found a similar distribution of results in my study to that in Bradlow, Ettliger, and Wong. There were groups that were either capable of learning the entire pattern (ceiling) or learned a regular, but simpler version of the pattern (conservative/anti-conservative participants, who either learned the static or dynamic mappings). When these participants were removed from the results, the dynamic results in all three conditions cluster around chance, as shown in table 6-25 (this table includes both multiclickers and non-multiclickers).

*Table 6-25: Scores with ceiling and strategic participants removed from the dataset*

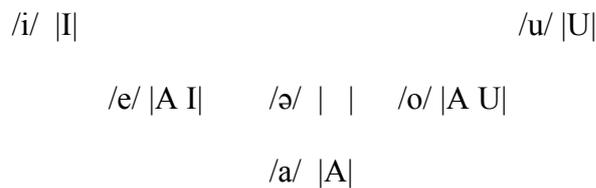
	Static	Dynamic	Total	<i>n</i>
Chain Shift	0.663	0.568	0.616	32
Coherent Raise	0.659	0.512	0.587	36
Mixed Raise	0.718	0.517	0.617	37

This suggests that, for these speakers, the pattern was not learned, which accords with Bradlow, Ettliger, and Wong’s third group of participants, who they call ‘non-learners’. Crucially, in none of the three groups does there appear to be a difference between the three conditions. The first thing that this strongly suggests is that if the study were to be re-run, using standardized tests of procedural and declarative

memory would be useful, in order to see if the similar distributions observed in the two studies have the same source (better declarative memory = better complex morphophonological pattern learning). A potentially novel finding of the present study is that, if participants who learn the whole pattern are using their skills of analogical reasoning, then the simplicity or otherwise of the pattern in terms of its phonology does not seem to matter.

For example, a prediction that was not borne out at all by the study is that the coherent condition should be the easiest to learn, given that unlike the other two conditions this condition can be generalized to one rule, [-high, -low] → [+high]. It is perhaps worth pointing out at this juncture that even if one does not subscribe to classical rule based phonology, it is still possible to describe the coherent raise condition in simpler terms than the other two conditions. Consider Element Theory (e.g., Harris & Lindsey (1995), Backley (2011)). In Element Theoretic terms, the vowel system present in the experiment can be defined as in figure 6-8.

*Figure 6-8: Element Theoretic treatment of the vowel system*



The coherent raise pattern, then, can be represented in terms of the loss of |A| from complex segments; a simple rule with one statement. The chain shift condition, on the other hand, must be described first as the addition of |I| to simplex |A|, but then the loss of A from complex |A I| but not complex |A U|. The mixed raise condition begins in the same way, but the second rule is the reverse of the second rule in the chain shift condition; |A| is deleted from underlying |A U| but not underlying |A I|. As we do not see a significant advantage for the coherent raise condition over the other two kinds of condition, it is not possible to claim that participants across conditions are employing a radically different strategy to deal with patterns based on one rule as opposed to patterns based on two rules. A potential reason for this is that, at least among speakers who learn the pattern, no rules are formed, as this is the work of procedural memory (see Ullman et al (1997)). In all three conditions, participants simply remember what changes and what does not. Given that the same

overall number of vowels change in each condition, we would predict that this kind of strategy should be equally easy across conditions.

The question that this leaves is whether this finding has any insight to offer in terms of genuine language acquisition. This is a concern that plagues all AGL literature, to the extent that Moreton & Pater begin their review of AGL studies by explicitly asking “do these highly artificial tasks reveal anything about natural language phonology?” (2012a, p.686). It is certainly true that the efficacy of AGLs has been questioned in the past. Redington & Chater (1996), discussing syntactic AGLs, state that even when AGLs appear to be successful (i.e., when participants appear to show an effect of learning), no genuinely abstract knowledge is required for this learning to take place. Whilst it is true that AGL studies on infants appear to show that the same kinds of biases that can be found by testing adults are also present in children (see Saffran & Thiessen (2003) or White & Sundara (2014) for examples of this), it is also true that there are studies whose authors make clear that the principles under discussion are not genuine principles of first language acquisition, but those of L2 learning. One such paper is Ettliger, Bradlow, and Wong (2014), in which the authors state that their findings “might account for individual variation in second language learning” (p.1).

Citing Williams (2005), Ettliger, Bradlow, and Wong refuse to explicitly generalize their results from L2 acquisition to L1, as “L2 is differentiated from L1 learning by its reliance on domain-general cognitive abilities” (p.18). As they point out, children generally have weaker declarative memory relative to procedural memory (as explored by Greenbaum & Graf (1989)). It is not clear that this method of acquisition would necessarily be available to infants. Additionally, if speakers are using domain-general pattern learning abilities, rather than any kind of implicit, phonological knowledge, then perhaps we would not expect to see any difference between the conditions. The three patterns in the experiment differ in terms of phonological complexity, but in a wider sense (how many sounds need to be learned? How many sounds are changing?), all three conditions are exactly the same. In all conditions, there are five vowels. Two out of those five vowels change. A potential strategy for learning the pattern would be to explicitly think of rules, of the sort ‘a becomes e, and e becomes i’. These rules do not require any complex

phonological knowledge. Indeed, it appears that this is what some participants did. In the third sub-study, I widened the post-test questionnaire to include the question ‘Did you use any kind of strategy during the experiment? If so, what was it?’ Answers to this question included those below:

“I was familiar with the concept having learned Spanish verb conjugations in the past, so that probably helped me pick up the patterns” (participant QGI145, overall score 100%)

“...So if you want to say that was a ‘strategy’ per se, I would say pattern recognition” (participant QDP147, overall score 100%)

“I tried to categorize all of the words using the vowel sounds at the end in order to remember them more easily” (participant BOR158, overall score 53.125%)

These kinds of responses suggest that the kind of pattern learning that was occurring was explicit. These participants (and others) realized that there was some change between the singular and plural and tried to work out what those changes were.

Whilst it is difficult to interpret all of the comments in terms of whether they represent a strategy or not as many of the comments were somewhat vague, it does appear that a clear majority were using some kind of strategy, though not always one that was germane to the task. However, some participants did suggest that they were using more implicit learning practices, as shown in the comments below:

“I can confirm that I didn’t use a particular strategy in learning the language, instead I just relied on answering from memory” (participant WAI145, overall score 100%)

“I didn’t employ any particular strategy in learning the language, as I wasn’t sure what types of question would be asked at the end. I just repeated each word aloud as instructed” (participant ZJX169, overall score 71.875%)

“No” (Participant QZY193, overall score 78.125%)

These qualitative results are inherently somewhat unsatisfying, and any subsequent study should definitely include a question on whether participants did in fact use some kind of strategy.<sup>45</sup> The fact that in the third sub-study some kind of strategizing

---

<sup>45</sup> The demographic questionnaire for all three sub-studies included the question “Can you guess the purpose of this experiment?” and invited participants to say what they thought it was if so. The breakdown of answers was 17.23% ‘Yes’ (31/181), 22.65% ‘Maybe’ (41/181), 56.90% ‘No’ (103/181), with 2.76% (5/181) participants not answering the question. Of those who said ‘Yes’ and ‘Maybe’ and offered some further comment, 18 (9.94% of the sample) gave an answer that specifically referred to changes in vowel sounds between singular and plural.

does seem to be prevalent suggests that more needed to be done in the training phase to provide distraction for participants, so that they would not have found it so easy to identify the pattern. This is not necessarily a huge problem. Moreton and Pertsova (2015) discuss AGL experiments with regard to whether they are able to tease out implicit or explicit learning strategies in participants, and they point out that both “[i]mplicit and explicit processes are available in phonological learning” (abstract).

The fact that learning is explicit does not necessarily make it totally non-phonological. Indeed, we might still expect some patterns to be easier than others to learn, even if that learning is explicit. Additionally, complex morphophonological processes like this are very difficult to teach implicitly. Whilst in some senses the experiment was quite explicit, certain steps were taken to lessen the explicitness of the task. According to Moreton & Pertsova, triggers for participants to begin to use explicit rule-learning systems include giving explicit feedback, which this experiment did not, and instructing participants to find a rule, which again, this experiment did not. Participants were simply instructed to listen to the words and repeat them back.

Even so, the feedback suggests that many participants were engaging their explicit, pattern-learning tendencies. This may have led them to ignore the phonological complexities inherent in the pattern in a way that perhaps language learning infants would not have done. Therefore, I am reticent to make any substantive claims about what the study says about phonological organization. However, what is clear is that the study does not offer any kind of support to the notion that the synchronic phonology is involved in the learning of processes like chain shifts which are purely morphologically triggered. The following section details potential improvements that could be made to following experiments in order to tease apart the results:

### *6.7 Further studies*

In future studies, there are a number of areas in which further control would be very useful. Perhaps most importantly, more information must be extracted from each participant. As mentioned above I would follow Bradlow, Ettliger and Wong (2014) in conducting tests on procedural, declarative and working memory, so that participants could be divided into groups based on their scores. This would enable

me to further examine the apparent correlation between good declarative memory and high achievement on the task, which would in turn allow for the stronger postulation that people who are able to learn the pattern are not using standard phonological rules, which would then allow for the stronger argument that the patterns should in fact, be just as easy as one another to learn, as they can only be learned by participants who are using an explicit learning strategy. As well as this, all participants in subsequent studies should be asked more detailed questions about any strategy that they used, and failure to answer these demographic questions should lead to the participant's data not being considered.

It is also possible that the reason that participants were able to ascertain what pattern was being presented and use their more explicit pattern-learning mechanisms was that, whilst the task itself was difficult, participants were too aware of what they were being asked to do. There are several ways in which the purposes of the task could be better concealed. For instance, the test phase could include questions about the semantic relationships between the pictures and sounds, forcing the participants to focus on an unrelated issue. Alternatively, a form of distraction completely divorced from the task could be used, such as asking maths or logic questions between training and test blocks. A third potential improvement could be a more explicit class of fillers that are nothing to do with the overall pattern. For example, another distinction, like gender, could be introduced, where the pattern again includes no vowel shift. Whilst there are many AGL studies that do include fillers that are completely disregarded in later analysis (see, e.g., Wilson (2006), White (2014)), there are others which do not feature these distractor stimuli (see, e.g., Pycha et al. (2003), Moreton (2008), Carpenter (2010), Finley & Badecker (2012), Ettliger, Bradlow & Wong (2014)). This is my least preferred option, since a real danger is biasing participants further towards conservatism by adding more stimuli that do not feature vowel shift. If such a class of fillers were added, then they would have to feature some kind of completely unrelated suppletive morphology.

A more radical departure from the current paradigm that would provide a far more implicit method of teaching would be to perform an EEG experiment. EEG experiments, in which brain wave activity is measured by the application of multiple electrodes to the scalp of a participant, are often entirely passive in their training

phase, with participants simply listening to stimuli whilst attending to a completely unrelated silent visual stimulus. These kinds of studies have been used successfully with the Artificial Grammar paradigm in the past. Attaheri et al (2015) succinctly summarize how the general procedure would work.

“Typically, in AG learning studies the participant is exposed to exemplary sequences of stimuli generated by the AG. In a subsequent testing phase, the participant’s response to ‘consistent’ sequences that follow the AG are evaluated relative to those that violate it. Different responses to ‘violation’ versus consistent sequences can provide evidence that the participant is sensitive to the properties of the AG” (p.74)

This is not an especially different methodology from that which is employed in the experiment above. There is a passive exposure phase, and a test phase where the differences in response to conforming and non-conforming stimuli are measured. The two main advantages to this kind of study are: 1) no instruction needs to be given, so the likelihood of participants engaging in explicit pattern-solving is significantly reduced; 2) Measuring differences in brain activity is a much more subtle kind of measurement than measuring responses to forced-choice questions, and less influenced by error, as the changes in brain activity are not conscious. I am not aware of any phonological AGL studies that use the EEG paradigm, but there is no *a priori* reason that this could not work. There are certainly many EEG studies that are phonological in nature, suggesting that there are relevant potentials that can be measured that pertain to phonological information (for studies see, for example, Eulitz & Lahiri (2004), Cornell, Lahiri & Eulitz (2011)).

Finally, I encountered certain practical problems running the experiment online as opposed to in laboratory conditions. The first and most significant of these, which I have already addressed at length, is the problem of multclicking. This is not the fault of the online nature of the study, but instead the lack of clear instructions, and so would be simple to fix in reruns of the experiment. However, there are certain aspects of the experiment that are impossible to control for if the experiment is online. For example, whilst explicit instructions were given to participants not to begin the task until volume was at a comfortable level, it is not possible to ensure

that participants adhered to this instruction. Equally, participants were instructed to be in a quiet room, and to use headphones. The wide variation in headphone quality, and the perception of what counts as a quiet room, means that it is impossible to be sure that all participants had the same experience of the experiment.

At this point, I do feel that a little defence of the decision to run the study online is in order. Even as a somewhat exploratory pilot study, I took data from 181 participants overall, (60 participants each in the chain shift and mixed raise conditions, and 61 in the coherent raise condition). This is a marked increase on the amount of participants used in most AGL studies. Table 6-26 below shows the participants used in four other AGL studies, all lab-based.

*Table 6-26: A comparison with previous studies*

Study	<i>n</i>
White (2014) – Study 1	40 (20 per condition)
Finley & Badecker (2012) – Study 1	60 (20 per condition)
Carpenter (2010) – Study 1	40 (21 in one condition, 19 in the other)
Moreton (2008) – Study 1	25 (1 condition, within-subjects design)

This shows that 20 participants per condition appears to be the norm in studies of this nature. Whilst I do believe that there would be many advantages in running a lab-based study, I also believe that it would have to be on a larger scale than the studies listed in table 6-26, given that the pattern appears to be learned differently, and in the majority of cases is harder to learn, than the patterns represented in the studies above.

### *6.8 Some speculations: What would it mean if the result were replicated?*

It should be clear from the foregoing that this experiment has certain methodological issues that suggest that its results do not present clear evidence that there is no difference in learnability between chain shift patterns and other, similar patterns. However, it is important to consider what it would mean if the results were replicable. This result would mean, in summary, that any pattern like chain shifting appears to only be learnable if participants use explicit learning strategies that use the broad, analogical methods associated with declarative memory, and if these methods are not available (potentially because participants do not perform well on

general tests of declarative memory), participants tend to perform at chance, meaning that the pattern is not learnable. This finding would accord with the wug-test studies presented by Nagle (2013; 2016) and Zhang, Lai & Sailor (2006) in their overall finding that chain shift patterns are not fully generalizable. Two further things would be suggested by a replication of the kind of AGL detailed above. The first is that this failure of generalizability would not be confined to one language. As I have pointed out, wug-testing for a property of one language is invaluable in the study of that language, but it tells us less about general phonological principles.

As well as this, it would suggest more generally that complex patterns (in the Bradlow, Ettliger and Wong sense of complex, i.e., no overt triggers) are *all* hard to learn for participants without particularly strong declarative memory. This would be a particularly interesting finding in the light of Bradlow, Ettliger, and Wong's discussion, where they speculate that part of the reason for the difficulty that participants had in learning the complex pattern was the opaque nature of the rule, particularly as (in their view at least, (2014, p.823)), English phonology does not contain opaque processes. If it were the case that the finding that there is no real difference between any of the conditions were replicated, this would suggest that the opaque nature of chain shifting as compared to the transparent nature of the other two processes is unimportant. It would simply be the case that most participants struggle to learn these kinds of complex morphophonological alternations, regardless of the genuine phonological complexities that separate the three conditions.

If further study cannot find evidence that participants use active synchronic phonology (of the kind that can be accessed by speakers with, for example, good procedural memory but comparatively poor declarative memory) when learning these kinds of patterns, then this would, eventually, constitute evidence against *any* system of phonology that had room for these kinds of processes in its architecture. Ullman et al (1997) state that "the lexicon is part of a... "declarative" memory system and... grammatical rules are processed by a ... "procedural" system" (p.266). This suggests that rather than the application of a grammatical, phonological rule, if participants are able to learn the patterns then what they are learning is a set of lexical items that can be used as the basis of analogical extension. This domain-

general explanation speaks against Rule-Based and Optimality Theoretic accounts of chain shifting, which describe such processes in terms of rules or constraints.

I am not stating that this study constitutes anything like solid evidence for the position that the information needed for chain shift effects resides purely in the lexicon and does not require any kind of phonological computation. Indeed, there may well be phonological computation involved in many of the processes that I have studied that have  $A \rightarrow B \rightarrow C$  orderings. In the previous five chapters I have taken the position that *some kind* of phonological computation is occurring, though I have taken issue with the notion that some specific mechanism must be built into the grammar to accommodate this. However, this study can be taken as a starting point, by adding another set of results to the (admittedly very small) pile of evidence which has, thus far, failed to find any evidence that speakers access their phonology at all when presented with either the extension of a chain shift pattern in their own language (as in Nagle (2013; 2016), Zhang, Lai and Sailor (2006)), or in the context of an artificial language.

## Chapter 7: Conclusions

### 7.1 Introduction

This thesis has argued that synchronic chain shifts do not exist in any interesting sense. Whilst I acknowledge that there are many phonological processes that result, or appear to result, in  $A \rightarrow B \rightarrow C$  mappings, I argue that this, in and of itself, is not especially important. The next section of this final chapter summarises my main findings chapter-by-chapter. I finish by briefly discussing the implications of these findings.

### 7.2 Main findings

The thesis embarked from the intuition that the general theories of phonological computation that have been used to discuss synchronic chain shifting (discussed in chapter 2) assume that the entire set of chain shifts, or at least a large proportion of this set, can be modelled in the same way. Because chain shifts are problematic for most theories of phonology, these theories all add substantial power to the phonological grammar in order to achieve this ‘one-size-fits-all’ solution. One example of the reach of theories like this is the assumption made by some theorists that there are substantive similarities between chain shifts in synchrony, and effects that evince a similar  $A \rightarrow B \rightarrow C$  mapping in diachrony and acquisition. In chapter 3, I argued that direct comparison of this kind is unhelpful. There are substantive differences between chain shifts in diachrony and synchrony, and it is not at all clear that child chain shifts in acquisition actually require the intervention of the phonological grammar.

Another example is the corpus of effects in Moreton (2004a), in which a set of hugely disparate effects is collected as a test of the principles of one particular chain shift theory in OT, Local Conjunction. In chapter 4, I performed a thorough reanalysis of the effects in this sample, as well as adding to the corpus with other examples of processes that have been called chain shifts in the recent phonological literature. I examined the trends in the chain shifts in the corpus, and noted that whilst there were certain kinds of shifts that do seem to re-occur time and again (for example, shifts based on vowel height or changes in the manner of consonants), over

half of the shifts in Moreton's corpus shared nothing with any other shift in the corpus except for their superficial  $A \rightarrow B \rightarrow C$  mappings. That is to say, the set of chain shifts contains an extremely wide variety of processes. There is no a priori reason to suggest that the same theoretical apparatus should be used to model all of these processes.

Chapter 5 is a more detailed investigation of certain particularly relevant shifts, which essentially amounts to a winnowing down of the corpus. The first part of the chapter discussed shifts that have been misdescribed, or are based on insufficient or confusing data. The next part of the chapter was an investigation of putative shifts that take place above the level of the segment. In the case studies I examine, it is difficult to justify a chain shift analysis of any pattern above the level of the segment. This is because the ordering effects either fall out naturally from more basic principles of the grammar, or because there is no  $A \rightarrow B \rightarrow C$  mapping present in the first place. This part of the chapter concludes by suggesting that the first step towards an accurate picture of chain shifting would be to use one of the more constrained definitions that has been suggested in the literature (see, e.g., McCarthy (2007), Baković (2011)), where chain shifting is seen as counterfeeding-on-*focus*, meaning that it only relates to effects that take place at the level of the segment.

The final part of chapter 5 investigated two of the most prominent kinds of effect that take place at the level of the segment: mutation and metaphony. I began by arguing that there are good reasons for considering mutation and metaphony to have certain substantive similarities. This may seem counterintuitive given that I have been arguing across the rest of the thesis that it is misguided to seek one overarching solution for all kinds of shifts. However, whilst I acknowledge that there are differences between consonant mutation and metaphony, I believe that the similarities they share are more important. Perhaps the most important of these similarities is that in both cases, the fact that we observe chain shift effects at some points in some individual dialects can be argued to be completely epiphenomenal. Seeming chain shift mappings are often parts of larger effects, even within the same language or dialect. In consonant mutation patterns, for example, it is possible for  $A \rightarrow B \rightarrow C$  mappings to coexist with exactly the kind of  $A \rightarrow C$  mappings that chain shift theories are designed to preclude.

Additionally, a cross-dialectal analysis of Italian metaphony suggests that any specific grammatical contrivance required to block an  $A \rightarrow C$  mapping is only necessary for dialects that have chain shift patterns. Because adding such a mechanism into the grammar is such a costly move, it is germane to ask whether it is plausible that learners would actually do this. There are ways of dealing with particular kinds of seeming chain shifts in metaphony and mutation that do not rely on a mechanism blocking  $A \rightarrow C$ . I argue that these changes operate on more general principles than the specific ad-hoc devices that have been suggested for dealing with chain shifts. I finish this chapter by pointing out that there are certain theorists who do not believe that metaphony and mutation involve any phonological computation at all, instead being purely morphological processes, thus instances of lexical selection rather than derivation. In a way, it is not especially important whether the computation that occurs in these kinds of effects is phonological or morphological, for our purposes. This is because, either way, I believe that it is inaccurate to call any of the processes that I have discussed in the corpus in chapter 4 a chain shift.

However, whether speakers are learning a genuinely phonological alternation (or set of alternations) when they learn a chain shift is an empirical question, and one with potentially interesting consequences. For this reason, chapter 6 sketches out a method that we could use to test this, and provides a pilot experiment showing how this sort of study might work. In this pilot Artificial Grammar Learning study, participants were presented with one of three artificial language fragments, each containing a different kind of vowel shift effect. Although these effects differed substantially in terms of featural complexity, the experiment failed to find any evidence that this had an influence on participants' performance.

There are two important caveats here. The first is that a negative result does not constitute evidence for the null hypothesis. In essence, it may be the case that there is a difference in learnability between chain shifts and other, minimally different patterns, but that this was the wrong type of experiment to properly elicit that difference, or that the difference was too subtle to be caught by the performance metrics used in the experiment. The second is that there were several methodological and design errors with the experiment that could potentially have obscured a result

that did feature a significant difference between the conditions. However, the experiment does present a first step towards the kind of study where we might ascertain exactly what learners are learning when they acquire complex morphophonological patterns (see also Bradlow, Ettliger & Wong (2014)).

Overall, the thesis suggests that synchronic chain shift is a label that is essentially empty, beyond a vague  $A \rightarrow B \rightarrow C$  mapping generalization. I also argue that this  $A \rightarrow B \rightarrow C$  mapping is found less often than is previously reported, and even when it does exist it is less fundamental to the nature of the effects under discussion than has been assumed in the majority of the literature. This relative lack of importance suggests that we, as linguists, should not be employing special grammatical mechanisms in order to create this  $A \rightarrow B \rightarrow C$  mapping, as this does not reflect a coherent set of generalizations that are salient to the speaker/hearer. In short, the general principles of phonology and/or morphology will give you an  $A \rightarrow B \rightarrow C$  mapping, or they won't.

### **7.3 Implications**

Synchronic chain shifts have received a great deal of attention in the phonological literature because they are generally seen as problematic under standard assumptions of what a phonological theory can do. This is particularly true of classical OT, which is completely incapable of modelling effects of this nature. Thus, chain shifts have been used to attempt to justify particular outgrowths of the basic architecture of the system (by, for example, Gnanadesikan (1997), Lubowicz (2003a), Moreton & Smolensky (2002) and Mortensen (2006)). If, as I argue, there is no coherent set of chain shift effects, then these theories begin to seem less necessary. It is important to reiterate at this stage that all of the theories that I discuss have a wider purview than chain shifting. My refutation of the concept of synchronic chain shifting does, therefore, not constitute a solid argument against the totality of these theories. However, it does suggest that all of these theories are capable of modelling something that, in all probability, they should not. In short, they are too powerful.

In order to test this intuition further, it would be instructive to examine the other kinds of effects that these specific theories attempt to model. For example, Padgett (2002, p.7) gives a useful list of kinds of phonological processes that Local

Conjunction analyses have been applied to, which is made up of coda neutralization to the unmarked, dissimilation, restrictions of triggers on assimilation, and derived environment effects. There is already a great deal of work on most, if not all, of these phenomena across the phonological literature, as is the case for chain shifting. A good first step, following the approach of this thesis, would be to perform a similar kind of gathering together of research on these processes, to investigate to what extent these processes actually represent coherent classes.

Another important line of enquiry that I have already discussed (see the final sections of chapter 6) is a broadening of experimental research into chain shifts. It is clear that the experiment presented here does not give a clear answer as to whether phonological knowledge genuinely plays a role in L1 learning of chain shift patterns. It is important, therefore, that a larger-scale, better controlled study be performed, in order that we may get more insight into whether it is plausible to suggest that the phonological grammar plays any role at all in chain shifting processes. I gave many suggestions for how this might be performed in the previous chapter (see especially sections 6.6 and 6.7).

#### **7.4 Concluding remarks**

This thesis has argued against the concept of synchronic chain shift. Comparison both with shifts in diachrony and acquisition and across the set of putative synchronic chain shifts shows that, aside from a superficial  $A \rightarrow B \rightarrow C$  mapping, there is very little of substance that unites most processes that have been called chain shifts in the literature. Additionally, the very notion of this  $A \rightarrow B \rightarrow C$  mapping is suspect in a surprising proportion of cases.

I believe that the findings of this thesis have certain implications for phonological theory. One important conclusion is that attempts to model synchronic chain shifts are reflections more of the knowledge that linguists have about languages than the knowledge speaker/hearers have of those languages. I have written in favour of a simpler model of phonological grammar that does not require a mechanism whose job is specifically to penalize  $A \rightarrow C$  mappings. I have also sketched out an experimental research programme that aims to provide evidence for whether chain shifts are amenable to phonological computation by speaker/hearers. Even if there

are effects featuring  $A \rightarrow B \rightarrow C$  mappings that do require genuine phonological computation, I argue that there are already-available ways of modelling these effects that do not rely on an  $A \rightarrow C$  blocking mechanism.

## Appendix A: A reanalysis of Mascaró's typology of Italian metaphony dialects

In attempting to show that the Preserve Contrast constraint is unnecessary in any non-chain shifting dialect, I limit myself to using only the other apparatus that Mascaró himself uses. That is to say, I assume the same featural suffix, Wolf's FLOAT-STRESS constraint, and general faithfulness constraints (DEP, ID(ATR) and ID(HIGH)).

In cases of neutralization, and cases where there is no change at all in low-mid vowels, contrast preservation constraints are unnecessary. Indeed, all that is required in terms of changing the analysis from Mascaró's original analysis is the removal of the preserve contrast constraint. Firstly, I show Mascaró's original tableaux for the kind of dialect that would feature neutralization (2015, pp.14-15), followed by the same tableau without the Preserve Contrast constraint:

*Table A-1: Mascaró's analysis of Netutralization*

/ε/	DEP-C	FLOAT-STRESS	PRCONT(STR, MID)	ID(hi)	ID(ATR)
ε		*!*			
e		*!			*
↻i			*	*	*
jε	*!				

*Table A-2: Neutralization analysis without contrast preservation constraint*

/ε/	DEP-C	FLOAT-STRESS	ID(hi)	ID(ATR)
ε		*!*		
e		*!		*
↻i			*	*
jε	*!			

It is clear from that tableaux above that the contrast preservation constraint is unnecessary. The only form that violates it is the winning candidate /i/, which means

that it must be low ranked enough that it does not figure in the choice of optimal candidate in a meaningful way.

The tableaux below illustrate how a dialect in which there is no change to the values of low-mid vowels can also be modelled in exactly the same way with or without the contrast preservation constraint:

*Table A-3: Mascaró's analysis of dialects with no low-mid metaphony*

/ε/	DEP-C	PRCont(STR, MID)	ID(ATR)	FLOAT-STRESS	ID(hi)
☞ ε				**	
e			*!	*	
i		*!	*		*
jε	*!				

*Table A-4: No low-mid metaphony analysis without contrast preservation constraint*

/ε/	DEP-C	ID(ATR)	FLOAT-STRESS	ID(hi)
☞ ε			**	
e		*!	*	
i		*!		*
jε	*!			

Mascaró uses the contrast preservation constraint to rule out [i]. However, it is not necessary to do this. An /ε/ → [i] mapping changes the specification of [ATR], so high-ranking ID(ATR) rules out both [e] and [i]. The cases where diphthongs are involved are more complex, as it seems at first blush that the contrast preservation constraint is required to explicitly block a mapping of /ε/ → [i]. It appears, then, that the constraint is doing the same work in dialects with diphthongization as it is in chain shift dialects, in that it is blocking an A → C mapping. Whether it needs to is a different question. I begin by discussing Mascaró's analysis of dialects where the diphthong that is created consists of a glide and a low-mid vowel. Table A-5 shows the analysis:

Table A-5: Mascaró's analysis of a dialect that creates low-mid diphthongs

/ε/	PRCONT(STR, MID)	FLOAT-STRESS	DEP-C	ID(hi)	ID(ATR)
ε		*!*			
e		*!			
i	*!			*	*
☞jε			*		

Outputs [ε] and [e] are ruled out by FLOAT-STRESS, whilst [i] is ruled out by PRCONT(STR, MID). It turns out that if Mascaró's constraints are re-ranked so that ID(ATR) outranks DEP-C, then the PRCONT constraint is unnecessary. Because I have introduced constraint re-ranking here, I show that this new ranking will work for the /ε/ → [i] part of the effect too:

Table A-6: Low-mid diphthong dialect analysis without contrast preservation constraint (input /ε/)

/ε/	FLOAT-STRESS	ID(ATR)	DEP-C	ID(hi)
ε	*!*			
e	*!	*		
i		*!		*
☞jε			*	

Table A-7: Low-mid diphthong dialect analysis without contrast preservation constraint (input /e/)

/e/	FLOAT-STRESS	ID(ATR)	DEP-C	ID(hi)
ε	*!*	*		
e	*!			
☞ i				*
jε		*!	*	

Mascaró claims that a re-ordering of his constraints can also provide an analysis of dialects like Calvello Italian, in which diphthongization leads to open vowels. However, it should be borne in mind that in terms of violation profile, there appears to be no difference, on Mascaró's account, between /jε/ and /je/. In the first four of

his five tableaux, only [jɛ] is evaluated as a candidate, and it violates only DEP-C. In his fifth tableau, by contrast, only [je] is evaluated, and again it violates only DEP-C. This suggests that further constraints would be needed to accurately model diphthongization effects, but this lies beyond the scope of this discussion.

Finally, we can consider the tableau for a chain shift dialect like Servigliano:

Table A-8: Mascaró's analysis of a chain shifting dialect

/ɛ/	DEP-C	PRCONT(STRESS, MID)	FLOAT-STRESS	ID(ATR)	ID(hi)
ɛ			**!		
e			*	*	
i		*!		*	*
jɛ	*!				

In this analysis, PRCONT rules out [i]. Given that [ɛ] is a [-ATR, -high] vowel, it does not satisfy FLOAT-STRESS on either count, and so violates it twice. On the other hand, underlying /e/ is [+ATR, -high], and so satisfies FLOAT-STRESS on one count whilst violating it on the other. Because of this, it wins out over [ɛ]. For input /ɛ/, ID(hi) could be re-ranked to replace PRCONT, because [i] violates ID(hi) whereas [e] does not. However, the ranking must work for both parts of the shift. For underlying /e/, a ranking of DEP-C >> ID(hi) >> FLOAT-STRESS >> ID(ATR) would not work, because both [e] and [ɛ] would win out over [i], which would again violate ID(hi).

## Appendix B: Stimuli for the AGL pilot study

Table B-1: Stem vowel /a/

Stem	Dual	Plural options
bətavə	bətavəsə	bətavətə bətevətə
dənatə	dənatəsə	dənatətə dənətətə
gəbarə	gəbarəsə	gəbarətə gəberətə
gəpakə	gəpakəsə	gəpekətə gəpakətə
kəsadə	kəsadəsə	kəsədətə kəsadətə
lədalə	lədaləsə	lədalətə lədelətə
ləmarə	ləmarəsə	ləmarətə ləmerətə
mədagə	mədagəsə	mədəgətə mədagətə
məkəpə	məkəpəsə	məkəpətə məkəpətə
nəkənə	nəkənəsə	nəkenətə nəkanətə
pələbə	pələbəsə	pələbətə pələbətə
pənəsə	pənəsəsə	pənəsətə pənesətə
rəvamə	rəvaməsə	rəvemətə rəvamətə
səravə	səravəsə	səravətə sərevətə
təsabə	təsabəsə	təsebətə təsabətə
vəgatə	vəgatəsə	vəgatətə vəgetətə

Table B-2: Stem vowel /e/

Stem	Dual	Plural options
bəpenə	bəpenəsə	bəpinətə bəpenətə
bərebə	bərebəsə	bəribətə bərebətə
dələnə	dələnəsə	dələnətə dəlinətə
gəsədə	gəsədəsə	gəsədətə gəsədətə
kəgedə	kəgedəsə	kəgedətə kəgidətə
ləlepə	ləlepəsə	ləlipətə ləlepətə
mətemə	məteməsə	mətemətə mətimətə
nəvetə	nəvetəsə	nəvitətə nəvetətə
pəgemə	pəgeməsə	pəgemətə pəgimətə
rəkekə	rəkekəsə	rəkekətə rəkikətə
sədərə	sədərəsə	sədərətə sədirətə
səmesə	səmesəsə	səmisətə səmesətə
təbeve	təbevəsə	təbivətə təbevətə
təpekə	təpekəsə	təpekətə təpikətə
vənegə	vənegəsə	vənigətə vənegətə
vərelə	vəreləsə	vərilətə vərelətə

Table B-3: Stem vowel /i/

Stem	Dual	Plural options
bəmikə	bəmikəsə	bəmikətə bəbukətə
dənivə	dənivəsə	dənivətə dəmuvətə
dəpilə	dəpiləsə	dəpulətə dəpilətə
gətirə	gətirəsə	gətirətə gəturətə
kənimə	kəniməsə	kənimətə kənumətə
kəsīgə	kəsīgəsə	kəsugətə kəsīgətə
ləgipə	ləgipəsə	ləgipətə ləgupətə
mətītə	mətītəsə	mətutətə mətītətə
nərisə	nərisəsə	nərusətə nərisətə
nəvinə	nəvinəsə	nəvunətə nəvinətə
pəbisə	pəbisəsə	pəbusətə pəbisətə
rəvigə	rəvigəsə	rəvugətə rəvigətə
səbilə	səbiləsə	səbilətə səbulətə
təlibə	təlibəsə	təlibətə təlubətə
vəkipə	vəkipəsə	vəkupətə vəkipətə

Table B-4: Stem vowel /o/

Stem	Dual	Plural options
bənomə	bənoməsə	bənomətə bənumətə
dəmosə	dəmosəsə	dəmusətə dəmosətə
dəvolə	dəvoləsə	dəvolətə dəvulətə
gəpodə	gəpodəsə	gəpodətə gəpudətə
kəbogə	kəbogəsə	kəbogətə kəbugətə
kəbopə	kəbopəsə	kəbupətə kəbopətə
lərogə	lərogəsə	lərugətə lərogətə
məvonə	məvonəsə	məvunətə məvonətə
nəlorə	nəlorəsə	nəlorətə nəlurətə
nətosə	nətosəsə	nətusətə nətosətə
pəkopə	pəkopəsə	pəkopətə pəkupətə
rədovə	rədovəsə	rəduvətə rədovətə
rəsobə	rəsobəsə	rəsubətə rəsobətə
sətotə	sətotəsə	sətotətə sətutətə
təmokə	təmokəsə	təmokətə təbukətə
vəgolə	vəgoləsə	vəgulətə vəgolətə

Table B-5: Stem vowel /u/

Stem	Dual	Plural options
bəbuwə	bəbuwəsə	bəbivətə_bəbuwətə
dənulə	dənuləsə	dənulətə_dənilətə
gəkunə	gəkunəsə	gəkinətə_gəkunətə
gəluwə	gəluwəsə	gəlivətə_gəluwətə
kəgupə	kəgupəsə	kəgipətə_kəgupətə
lədubə	lədubəsə	lədubətə_lədibətə
ləkugə	ləkugəsə	ləkugətə_ləkiɡətə
mədukə	mədukəsə	mədukətə_mədikətə
məvubə	məvubəsə	məvubətə_məvibətə
nəsutə	nəsutəsə	nəsitətə_nəsutətə
pərusə	pərusəsə	pərusətə_pərisətə
pəturə	pəturəsə	pəturətə_pətirətə
rəpurə	rəpurəsə	rəpirətə_rəpurətə
sənutə	sənutəsə	sənitətə_sənutətə
təsumə	təsuməsə	təsımətə_təsumətə
vəmudə	vəmudəsə	vəmudətə_vəmidətə



## References

- Acharya, J. (1991). *A Descriptive Grammar of Nepali and an Analyzed Corpus*. Georgetown: Georgetown University Press.
- Aitchison, J. (2013). *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.
- Almihmadi, M. (2011). *Understanding the Phonetics of Neutralization: A Variability-Field Account of Vowel/Zero Alternations in a Hijazi Dialect of Arabic*. PhD Dissertation: UCL.
- Al-Mozainy, H. (1981). *Vowel Alternations in a Bedouin Hijazi Arabic Dialect: Abstractness and Stress*. PhD Dissertation: University of Texas at Austin.
- Andersen, T. (1993). Vowel quality alternation in Dinka verb inflection. *Phonology*, 10(1), 1-42.
- Anderson, B. (1999). /ai/ monophthongization among African-American Detroiters: Another case of dialect levelling. *Paper given at NWAV(E) 28*. Toronto.
- Anderson, B. (2002). Dialect levelling and /ai/ monophthongization among African-American Detroiters. *Journal of Sociolinguistics*, 6(1), 86-98.
- Anderson, S. (1978). Tone features. In V. Fromkin (Ed.), *Tone: A Linguistic Survey* (pp. 133-176). New York: Academic Publishers.
- Anttila, A. (2000). *Morphologically-Conditioned Phonological Alternations*. Boston University and National University of Singapore: ms.
- Archangeli, D., & Pulleyblank, D. (1994). *Grounded Phonology*. Cambridge, MA: MIT Press.
- Archangeli, D., & Suzuki, K. (1997). The Yokuts challenge. In *Derivations and Constraints in Phonology* (pp. 197-226). Cambridge: Cambridge University Press.
- Árnason, K. (1988). Problems in the lexical phonology of Icelandic. In W. Dressler, H. Luschutzky, O. Pfeiffer, & J. Rennison (Ed.), *Phonologica: Proceedings of the 6th International Phonology Conference* (pp. 5-14). Cambridge: Cambridge University Press.
- Árnason, K. (2011). *The Phonology of Icelandic and Faroese*. Oxford: Oxford University Press.
- Attaheri, A., Kikuchi, Y., Milne, A., Wilson, B., Alter, K., & Petkov, C. (2015). EEG potentials associated with Artificial Grammar Learning in the primate brain. *Brain and Language*, 148, 74-80.
- Backley, P. (2011). *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.
- Baer-Henney, D., & van de Vijver, R. (2012). On the role of substance, locality and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology*, 3(2), 221-249.
- Baković, E. (1999). *Assimilation to the Unmarked*. Rutgers Optimality Archive: Ms.
- Baković, E. (2007). A revised typology of opaque generalizations. *Phonology*, 24(2), 217-259.
- Baković, E. (2011). Opacity and ordering. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The Handbook of Phonological Theory*. Oxford: Wiley-Blackwell.
- Baković, E. (2013). *Blocking and Complementarity in Phonological Theory*. London: Equinox.
- Bandhu, C., & Dahal, B. M. (1971). *Nepali Segmental Phonology*. Kirtipur: Summer Institute of Linguistics.

- Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.
- Barrie, M. (2006). Tone circles and contrast preservation. *Linguistic Inquiry*, 37(1), 131-141.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Bauer, L. (1979). The second great vowel shift? *Journal of the International Phonetic Association*, 9(2), 57-66.
- Becker, M., & Levine, J. (2010). Experigen: An Online Experiment Platform.
- Becker, M., Nevins, A., & Levine, J. (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language*, 88(2), 231-268.
- Bermúdez-Otero, R. (2011). Cyclicity. In M. van Oostendorp, K. Rice, C. Ewen, & E. Hume (Eds.), *The Blackwell Companion to Phonology* (pp. 2019-2048). Oxford: Wiley-Blackwell.
- Blaho, S. (2008). *The Syntax of Phonology: A Radically Substance-Free Approach*. PhD Dissertation: University of Tromsø.
- Bleile, K. (2004). *Manual of Articulation and Phonological Disorders: Infancy Through Adulthood*. San Diego: Singular Publishing Group.
- Bloomfield, L. (1944). Secondary and tertiary responses to language. *Language*, 20(2), 45-55.
- Boersma, P. (1999). The Phonology/Semantics Interaction in OT, and its Acquisition. Rutgers Optimality Archive, ms.
- Boyle, J. (2002). Incorporation vs. compounding in Hidatsa relative clauses or an "incorporation" relative clause in Hidatsa that doesn't. *Proceedings from the Main Session of the Thirty-Eighth Meeting of the Chicago Linguistic Society*, 38, 103-112.
- Boyle, J. (2007). *Hidatsa Morpho-Syntax and Clause Structure*. PhD Dissertation: University of Chicago.
- Boyle, J. (No Year). *Hidatsa lessons*. Retrieved from [http://www.neiu.edu/~linguist/hidatsa\\_lessons\\_vocab1.html](http://www.neiu.edu/~linguist/hidatsa_lessons_vocab1.html)
- Bradshaw, M. (1996). One-step raising in Gbanu. *Ohio State Working Papers in Linguistics*, 48, 1-11.
- Bright, W. (1957). *The Karok Language*. Berkeley: University of California Press.
- Bromberger, S., & Halle, M. (1989). Why phonology is different. *Linguistic Inquiry*, 20, 51-70.
- Broś, K. (2015). *Survival of the Fittest: Fricative Lenition in English and Spanish from the Perspective of Optimality Theory*. Cambridge: Cambridge Scholars Publishing.
- Broselow, E. (2008). Stress-Epenthesis interactions. In B. Vaux, & A. Nevins (Eds.), *Rules, Constraints and Phonological Constraints* (Ch. 8). Oxford: Oxford University Press.
- Bybee, J., & Slobin, D. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 58(2), 265-289.
- Bye, P., & Svenonius, P. (2012). Nonconcatenative morphology as epiphenomenon. In J. Trommer (Ed.), *The Morphology and Phonology of Exponence* (pp. 427-495). Oxford: Oxford University Press.
- Calabrese, A. (1985). Metaphony in Salentino. *Rivista di Grammatica Generativa*, 9-10, 1-140.

- Calabrese, A. (1995). A constraint based theory of phonological markedness and simplification procedures. *Linguistic Inquiry*, 373-463.
- Calabrese, A. (1998). Metaphony revisited. *Rivista di Linguistica*, 7-68.
- Calabrese, A. (2005). *Markedness and Economy in a Derivational Model of Phonology*. Berlin: Mouton de Gruyter.
- Calabrese, A. (2011). Metaphony in Romance. In M. van Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *Blackwell Companion to Phonology*. Oxford: Wiley-Blackwell.
- Cammenga, J. (1994). *Kuria Phonology and Morphology*. PhD Dissertation: Free University, Amsterdam.
- Campbell, L. (1985). *The Pipil Language of El Salvador*. Berlin: Mouton de Gruyter.
- Campbell, L. (2004). *Historical Linguistics* (3rd ed.). Edinburgh: Edinburgh University Press.
- Carpenter, A. (2010). A naturalness bias in learning stress. *Phonology*, 27(3), 345-392.
- Chacha, C. N., & Odden, D. (1998). The phonology of vocalic height in Kikuria. *Studies in African Linguistics*, 27, 129-158.
- Chapallaz, M. (1979). *The Pronunciation of Italian: A Practical Introduction*. London: Bell & Hyman.
- Chen, M. Y. (1987). The syntax of Xiamen tone sandhi. *Phonology*, 4, 109-149.
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Clements, G. N. (1991). Vowel height in the Bantu languages. *Proceedings of the Seventeenth Meeting of the Berkeley Linguistics Society*, 17, 25-64.
- Clopper, C. (2001). *The Nzema verbal phrase: An Optimality Theoretic account*. Retrieved from Indiana University Linguistics Club Working Papers Online 1(4): <http://www.indiana.edu/~iulcwp>.
- Cohn, A. (1993). Nasalization in English: phonology or phonetics? *Phonology*, 10(1), 43-81.
- Cole, J. (1998). Deconstructing metaphony. *Rivista di Linguistica*, 10, 69-98.
- Contreras, H. (1969). Simplicity, descriptive adequacy, and binary features. *Language*, 45(1), 1-8.
- Corbett, G. (2005). The canonical approach in typology. In Z. Frajzyngier, A. Hodges, & D. Rood (Eds.), *Linguistic Diversity and Language Theories* (pp. 25-50). Amsterdam: John Benjamins.
- Cornell, S., Lahiri, A., & Eulitz, C. (2011). What you encode is not necessarily what you store. *Brain Research*, 1394, 79-89.
- Crowhurst, M. (2011). Constraint conjunction. In M. van Oostendorp, C. Ewen, B. Hume, & K. Rice (Eds.), *Companion to Phonology* (p. Ch.65). Oxford: Wiley-Blackwell.
- Crowhurst, M., & Hewitt, M. (1997). *Boolean Operations and Constraint Interactions in Optimality Theory*. UNC Chapel Hill and Brandeis University: Ms.
- D'Arcy, A. (2003). Yowlumne re-examined: A challenge for contrastive specification. *Toronto Working Papers in Linguistics*, 20, 21-46.
- Dasgupta, P. (1982). Phonology and the Bangla verb. *Indian Linguistics*, 43(1-2), 17-28.
- Dayley, J. (1985). *Tzutujil Grammar*. Berkeley: University of California Press.
- Dayley, J. (1989). *Tümpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.

- de Lacy, P. (2002). *The Formal Expression of Markedness*. PhD Dissertation: U-Mass, Amherst.
- de Lacy, P. (2006). *Markedness: Reduction and Preservation in Phonology*. Cambridge: University of Cambridge.
- de Rijk, R. (1970). Vowel interaction in Bizcayan Basque. *Fontes Linguae Vasconum: Studia et Documenta*, 2(5), 149-168.
- de Saussure, F. (1916/2011). *Course in General Linguistics*. New York: Columbia University Press.
- Dejna, K. (1993). *Dialekty Polskie*. Wrocław: Ossolineum.
- Dinnsen, D., & Barlow, J. (1998). On the characterization of a chain shift in normal and delayed acquisition. *Journal of Child Language*, 25, 61-94.
- Dinnsen, D., & Eckman, F. (1977). Some substantive universals in atomic phonology. *Lingua*, 45, 1-14.
- Dinnsen, D., Green, C., Gierut, J., & Morrisette, M. (2011). On the anatomy of a chain shift. *Journal of Linguistics*, 47(2), 275-299.
- Donegan, P., & Stampe, D. (1979). The study of Natural Phonology. In D. Dinnsen (Ed.), *Current Approaches to Phonological Theory (Conference on the Differentiation of Current Phonological Theories, Bloomington)* (pp. 126-173). Bloomington: Indiana University Press.
- Ettlinger, M. (2007). Variation as a window into opacity. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 43, pp. 61-75.
- Ettlinger, M. (2008). *Input-Driven Opacity*. PhD Dissertation: University of California, Berkeley.
- Ettlinger, M. (2009). Phonological chain shifts during acquisition: Evidence for lexicon optimization. *Proceedings of the 38th Annual Meeting of the Northeast Linguistics Society*, 38, 259-269.
- Ettlinger, M., Bradlow, A., & Wong, P. (2014). Variability in the learning of complex morphophonology. *Applied Psycholinguistics*, 35(4), 807-831.
- Eulitz, C., & Lahiri, A. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience*, 16(4), 577-583.
- Finley, S. (2009). Morphemic harmony as featural correspondence. *Lingua*, 119(3), 478-501.
- Finley, S., & Badecker, W. (2012). Learning biases for vowel height harmony. *Journal of Cognitive Science*, 13, 287-327.
- Flemming, E. (1995). *Auditory representations in phonology*. PhD Dissertation: UCLA.
- Flemming, E. (1995). Evidence for constraints on contrast. *Paper Presented at WCCFL 14, University of California, Los Angeles* (March 1995).
- Flora, M. (1974). *Palauan Phonology and Morphology*. PhD Dissertation: USC San Diego.
- Foley, J. (1970). Phonological distinctive features. *Folia Linguistica*, 4(1-2), 87-92.
- Ford, A., Singh, R., & Martohardjono, G. (1997). *Pace Panini: Towards a Word-Based Theory of Morphology*. New York: Peter Lang.
- Foster, M. L. (1968). *The Tarascan Language*. Berkeley: University of California Press.
- Frajzyngier, Z. (1985). *A Pero-English and English-Pero Vocabulary*. Berlin: D.Reimer.
- Frajzyngier, Z. (1989). *A Grammar of Pero*. Berlin: D.Reimer.

- Fukazawa, H., & Miglio, V. (1998). Restricting conjunction to constraint families. (V. Samian, Ed.) *Proceedings of WCCFL*, 9, 102-117.
- Fulcrand, J. (2015). A reanalysis of the Great English Vowel Shift under Contrast Preservation theory. *Linguistic Research*, 32(3), 533-571.
- Gamble, G. (1978). *Wikchamni Grammar*. Berkeley: University of California Press.
- Ghosh, T. (2001). Vowel harmony in Bangla: An Optimality account. In A. Abbi, R. Gupta, & A. Kidwai (Eds.), *Linguistic Structure and Language Dynamics in South Asia: Papers from the Proceedings of SALA XVIII Roundtable* (pp. 144-163). New Delhi: Motilal Banarsidass.
- Gibbon, F. (1999). Undifferentiated lingual gestures in children with articulatory/phonological disorders. *Journal of Speech, Language and Hearing Research*, 42, 382-397.
- Gibson, C. S., & Ringen, C. (2000). Icelandic umlaut in Optimality Theory. *Nordic Journal of Linguistics*, 23, 49-64.
- Gioscio, J. (1985). *Il dialetto Lucano di Calvello*. Stuttgart: Steiner.
- Gnanadesikan, A. (1997). *Phonology With Ternary Scales*. PhD Dissertation: U-Mass Amherst.
- Goldsmith, J. (1976) An overview of Autosegmental Phonology. *Linguistic Analysis*, 2, 23-68
- Gonda, J. (1966). *A Concise Elementary Grammar of the Sanskrit Language (translated by Gordon Ford)*. University, AL: University of Alabama Press.
- Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A., & Trudgill, P. (2004). *New Zealand English: Its Origins and Evolution*. Cambridge: Cambridge University Press.
- Gouskova, M., & Becker, M. (2013). Nonce words show that Russian -yer alternations are governed by the grammar. *Natural Language and Linguistic Theory*, 31(3), 735-765.
- Gouskova, M., & Hall, N. (2009). Acoustics of epenthetic vowels in Lebanese Arabic. In S. Parker (Ed.), *Phonological Argumentation: Essays on Evidence and Motivation* (pp. 203-225). London: Equinox.
- Green, A. (2006). The independence of phonology and morphology: The Celtic mutations. *Phonology*, 116, 1946-1985.
- Greenbaum, J., & Graf, P. (1989). Preschool period development of implicit and explicit remembering. *Bulletin of the Psychonomic Society*, 27, 417-420.
- Grijzenhout, J. (1995). *Irish Consonant Mutation and Phonological Theory*. OTS Dissertation: Utrecht.
- Grimaldi, M., Calabrese, A., Zagona, F., Garrappa, L., & Sisinni, B. (2010). Articulatory grounding of Southern Salentino harmony processes. *Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech)* (pp. 1561-1564). Makuhari, Japan: ISCA.
- Gurevich, N. (2004). *Lenition and Contrast: The Functional Consequences of Certain Phonetically Conditioned Sound Changes*. London: Routledge.
- Gussenhoven, C., & Jacobs, H. (2011). *Understanding Phonology* (3rd ed.). London: Hodder Arnold.
- Guthrie, M. (1968). Notes on Nzebi (Gabon). *Journal of African Languages*, 7, 101-129.
- Hale, M., & Reiss, C. (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, 656-685.

- Hale, M., & Reiss, C. (2008). *The Phonological Enterprise*. Oxford: Oxford University Press.
- Halle, M., & Clements, N. (1983). *Problem Book in Phonology: A Workbook for Introductory Courses in Linguistics and in Modern Phonology*. Cambridge, MA: MIT Press.
- Hansson, G. (2008). Diachronic explanations of sound patterns. *Language And Linguistics Compass*, 2(5), 859-893.
- Harries, L. (1950). *A Grammar of Mwera*. Johannesburg: Witwatersrand University Press.
- Harrikari, H. (2000). *Segmental Length in Finnish - Studies Within a Constraint-Based Approach*. PhD Dissertation: University of Helsinki.
- Harris, A. (2009) Exuberant Exponence in Batsbi *Natural Language and Linguistic Theory* Vol. 27 Iss. 2 pp.267-303
- Harris, J. (1990). Segmental complexity and phonological government. *Phonology*, 7(1), 255-300.
- Harris, J. (2006). The phonology of being understood: Further arguments against sonority. *Lingua*, 116(10), 1483-1494.
- Harris, J., & Lindsey, G. (1995). The elements of phonological representation. In F. Katamba (Ed.), *Frontiers of Phonology: Atoms, Structures, Derivations* (pp. 34-79). Harlow, Essex: Longman.
- Harris, Z. (1942). Morpheme alternants in linguistic analysis. *Language*, 18(3), 169-180.
- Hart, M. (2010). *Consonant Lenition in Danish*. MA Dissertation: University of Tromsø.
- Hayes, B. (1986). Assimilation as spreading in Toba Batak. *Linguistic Inquiry*, 17(3), 467-499.
- Hayes, B. (2009). *Introductory Phonology*. Oxford: Wiley.
- Heydorn, R. (1943-44). Das Manyá: Ein in Liberia gesprochener Mandingodialekt. *Zeitschrift für Eingeborenen-Sprachen*, 34, 25-53.
- Holt, D. E. (2003). *Optimality Theory and Language Change*. New York: Springer.
- Honeybone, P. (2001/2002). *Germanic Obstruent Lenition: Some Mutual Implications of Theoretical and Historical Phonology*. PhD Dissertation: University of Newcastle upon Tyne.
- Hopkins, P. (2001). *The Phonological Structure of the Kashubian Word*. PhD Dissertation: University of Victoria.
- Hsieh, F.-f. (2005). Tonal chain-shifts as anti-neutralization induced tone sandhi. *Proceedings of the 28th Penn Linguistics Colloquium, Penn Working Papers in Linguistics*, 11(1), 99-112.
- Hsieh, H.-I. (1976). On the unreality of some phonological rules. *Lingua*, 38, 1-19.
- Hualde, J. I. (1989). Autosegmental and metrical spreading in the vowel harmony systems of northern Spain. *Linguistics*, 27, 773-805.
- Hualde, J. I. (1991). *Basque Phonology*. London: Routledge.
- Ito, J., & Mester, A. (1997). Correspondence and compositionality: The ga-gyo variation in Japanese phonology. In I. Roca (Ed.), *Derivations and Constraints in Phonology* (pp. 419-462). Oxford: Oxford University Press.
- Jaeger, J. (1984) Assessing the Psychological Status of the Vowel Shift Rule *Journal of Psycholinguistic Research* Vol. 13 No. 1 pp.13-36
- Jakobson, R. (1928). The concept of the sound law and the teleological criterion. *Casopis pro Moderni Filologii*, 1-2.
- Jakobson, R., & Halle, M. (1956). *Fundamentals of Language*. The Hague: Mouton.

- Jensen, S. (1994). Is  $\text{ʔ}$  an element? Towards a non-segmental phonology. *SOAS Working Papers in Linguistics*, 4, 71-78.
- Jesney, K. (2005). *Chain Shift in Phonological Acquisition*. MA Dissertation: University of Calgary.
- Jesney, K. (2007). Child chain shifts as faithfulness to input prominence. In A. Belikova, L. Meroni, & M. Umeda (Ed.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America*, (pp. 188-199).
- Jespersen, O. (1909). *A Modern English Grammar on Historical Principles*. Copenhagen: Munksgaard.
- Josephs, L. (1975). *Palauan Reference Grammar*. Honolulu: University Press of Hawaii.
- Kaplan, L. (1981). Phonological Issues in North Alaskan Inupiaq. Fairbanks: University of Alaska.
- Kawahara, S. (2002). *Similarity Among Variants: Output-Variant Correspondence*. Undergraduate Dissertation: International Christian University, Tokyo.
- Kaye, J. (1974). Opacity and recoverability of phonology. *Canadian Journal of Linguistics*, 19, 134-149.
- Kaye, J., & Pochtrager, M. (2013). GP 2.0. *SOAS Working Papers in Linguistics*, 16.
- Kaye, J., & Pochtrager, M. (2014). Phony Metaphony. *Talk given at the 22nd Manchester Phonology Meeting*.
- Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990). Constituent structure and government in phonology. *Phonology*, 7, 193-231.
- Kaze, J. (1989). *Metaphony in Italian and Spanish dialects revisited*. PhD Dissertation: University of Illinois, Champaign.
- Keating, P. (1988). Underspecification in phonetics. *Phonology*, 5, 275-292.
- Keating, P. (1990). The window model of coarticulation: Articulatory evidence. In J. Kingston, & M. Beckman (Eds.), *Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech* (pp. 451-470). Cambridge: Cambridge University Press.
- Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.
- Kenstowicz, M., & Kisseberth, C. (1977). *Topics in Phonological Theory*. New York: Academic Press.
- Kenstowicz, M., & Kisseberth, C. (1979). *Generative Phonology*. San Diego: Academic Press.
- Kiparsky, P. (1970/1982). Historical Linguistics. In P. Kiparsky, *Explanation in Phonology* (pp. 45-57). Dordrecht: Foris Publications.
- Kiparsky, P. (1971/1982). Historical Linguistics. In P. Kiparsky, *Explanation in Phonology* (pp. 57-81). Dordrecht: Foris Publications.
- Kiparsky, P. (1973a). Abstractness, opacity and global rules. In O. Fujimura (Ed.), *Three Dimensions of Linguistic Theory* (pp. 57-86). Tokyo: TEC.
- Kiparsky, P. (1973b). Elsewhere in phonology. In P. Kiparsky, & S. Anderson (Eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston.
- Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology*, 2, 83-138.
- Kiparsky, P. (2000). Opacity and cyclicity. *The Linguistic Review*, 351-367.
- Kiparsky, P. (2015). Stratal OT: A synopsis and FAQs. In Y. Hsiao, & L.-H. Wee (Eds.), *Capturing Phonological Shades*. Cambridge: Cambridge Scholars Publishing.

- Kirchner, R. (1995). Going the distance: Synchronic chain shifts in Optimality Theory. Ms, Rutgers Optimality Archive.
- Kirchner, R. (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry*, 27(2), 341-350.
- Kirchner, R. (1998). *An Effort-Based Approach to Consonant Lenition*. PhD Dissertation: UCLA.
- Krämer, M. (2009). *The Phonology of Italian*. Oxford: Oxford University Press.
- Krämer, M. (2012). *Underlying Representations*. Cambridge: Cambridge University Press.
- Labov, W. (1994). *Principles of Linguistic Change, Volume 1: Internal Factors*. Oxford: Blackwell.
- Labov, W., Yaeger, M., & Steiner, R. (1972). *A Quantitative Study of Sound Change in Progress: Volume 1. Report on National Science Foundation Contract NSF-GS-3287*. Philadelphia: University of Philadelphia Press.
- Labrone, L. (2012). *The Phonology of Japanese*. Oxford: Oxford University Press.
- Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Lahiri, A. (2003). Hierarchical restructuring in the creation of verbal morphology in Bengali and Germanic: Evidence from phonology. In A. Lahiri (Ed.), *Analogy, Levelling, and Markedness*. Berlin: Mouton de Gruyter.
- Langstrof, C. (2006). Acoustic evidence for a push-chain shift in the intermediate period of New Zealand English. *Language Variation and Change*, 18(2), 141-164.
- Lass, R. (1984). *Phonology: An Introduction to Basic Concepts*. Cambridge: Cambridge University Press.
- Lass, R. (1997). *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- Lass, R. (1999). Phonology and Morphology. In *The Cambridge History of the English Language: Volume III (1476-1776)* (pp. 56-186). Cambridge: Cambridge University Press.
- Lavoie, L. (1996). Consonant strength: Results of a data base development project. *Working Papers of the Cornell Phonetics Laboratory*, 11, 269-316.
- Levi, S. (2008). Phonemic vs. derived glides. *Lingua*, 12, 1956-1978.
- Lieber, R. (1987). *An Integrated Theory of Autosegmental Processes*. Albany, NY: State University of New York Press.
- Lindau, M. (1978). Vowel features. *Language*, 54(3), 541-563.
- Lindsey, G. & P. Szigetvári (2016) *CUBE: Current British English Searchable Transcriptions* Available online at: <http://seas3.elte.hu/cube/>
- Linebaugh, G. (2007). *Phonetic Grounding and Phonology: Vowel Backness Harmony and Vowel Height Harmony*. PhD Dissertation: University of Illinois at Urbana-Champaign.
- Lorentz, F. (1958). *Gramatyka Pomorska*. Wrocław: Ossolineum.
- Łubowicz, A. (2002). Derived environment effects in Optimality Theory. *Lingua*, 112, 243-280.
- Łubowicz, A. (2003a). *Contrast Preservation in Phonological Mappings*. PhD Dissertation: University of Massachusetts, Amherst.
- Łubowicz, A. (2003b). Counter-feeding opacity as a chain shift effect. (G. Garding, & M. Tsujimura, Eds.) *Proceedings of WCCFL*, 22, 315-327.
- Łubowicz, A. (2005). Locality of conjunction. (J. Alderete, C. Han, & A. Kochetov, Eds.) *Proceedings of WCCFL*, 24, 254-262.

- Łubowicz, A. (2011). Chain Shifts. In M. van Oostendorp, C. Ewen, B. Hume, & K. Rice (Eds.), *Companion to Phonology*. Oxford: Blackwell.
- Łubowicz, A. (2012). *The Phonology of Contrast*. Equinox: London.
- Luick, K. (1896). *Untersuchungen zur Englischen Lautgeschichte*. Strasbourg: Trubner.
- MacKay, C. (1994). A sketch of Misantla Totonac phonology. *International Journal of American Linguistics*, 60, 369-419.
- Macken, M. (1980). The child's lexical representation: The puzzle-puddle-pickle evidence. *Journal of Linguistics*, 16, 1-17.
- Mahanta, S. (2007). *Directionality and Locality in Vowel Harmony*. PhD Dissertation: Utrecht University.
- Mahanta, S. (2012). Chain shifts, feature distance, and contrast preservation. *The EFL Journal (Special Issue on Optimality Theory)*, 3(2), 41-64.
- Maiden, M. (1991). *Interactive Morphology: Metaphony in Italy*. London: Routledge.
- Major, R. (2005). Chemehuevi revisited. *Journal of the Southwest*, 47(3), 523-533.
- Manuel, S., Shattuck-Hufnagel, S., Huffman, M., Stevens, K., Carlson, R., & Hunnicut, S. (1992). Studies of vowel and consonant reduction. *Proceedings of the 1992 International Conference on Spoken Language Processing*, (pp. 943-946).
- Marshall, M. (1984). *The Dialect of Notre-Dame-de-Sanilhac: A Natural Language Generative Phonology*. Farringdon, Oxfordshire: Anna Libri.
- Martinet, A. (1955). *Economie des Changements Phonétiques: Traité de Phonologie Diachronique*. Bern: Francke.
- Martínez-Gil, F. (2006). Upstepping vowel height: A constraint-based account of metaphony in Proto-Spanish and Lena Asturian. In F. Martínez-Gil, & S. Collina (Eds.), *Optimality-Theoretic Studies in Spanish Phonology* (pp. 99-145). Amsterdam: John Benjamins.
- Mascaró, J. (1976). *Catalan Phonology and the Phonological Cycle*. PhD Dissertation, MIT.
- Mascaró, J. (2011). An analysis of stress-dependent harmony in Servigliano. *Probus*, 23, 21-55.
- Mascaró, J. (2015). *On the Typology of Metaphony/Stress Dependent Harmony*. Retrieved from Joan Mascaró's Personal Website: <http://mascaro-uab.weebly.com/uploads/4/0/9/2/40925737/typolmetaph-press.pdf>
- Massam, D. (1983). The morphology of Irish mutation. *MIT Working Papers in Linguistics*, 5, 10-29.
- McCarthy, J. (1993). A case of surface constraint violation. *Canadian Journal of Linguistics*, 38, 169-196.
- McCarthy, J. (1999). Sympathy and phonological opacity. *Phonology*, 16, 331-399.
- McCarthy, J. (2002). Comparative markedness. In A. Carpenter, A. Coetzee, & P. de Lacy (Eds.), *Papers in Optimality Theory II: University of Massachusetts Occasional Papers in Linguistics 26* (pp. 171-246). Amherst, MA: GLSA.
- McCarthy, J. (2005). Taking a free-ride in morphophonemic learning. *Catalan Journal of Linguistics*, 19-55.
- McCarthy, J. (2007). *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.
- McElree, B., Jia, G., & Litvak, A. (2000). The time course of conceptual processing in bilingual populations. *Journal of Memory and Language*, 42, 229-254.

- Menn, L., & Stoel-Gammon, C. (1993). Phonological development. In J. Berko Gleason (Ed.), *Language Development* (pp. 65-113). Columbus: Merrill.
- Mielke, J., Armstrong, M., & Hume, E. (2003). Looking through opacity. *Theoretical Linguistics*, 29(1-2), 123-139.
- Miglio, V., & Morén, B. (2003). Merger avoidance and lexical reconstruction. In E. Holt, *Optimality Theory and Language Change* (pp. 191-228). Kluwer: Dordrecht.
- Miller, G., & Chomsky, N. (1963). Finitary models of language users. In R. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology, Vol 11* (pp. 419-492). New York: Wiley.
- Mills, E. (1984). *Senoufo Phonology, Discourse to Syllable (a Prosodic Approach)*. Dallas, TX: Summer Institute of Linguistics.
- Mohanan, K. (1986). *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Moreton, E. (2004a). *Synchronic Chain Shift (Counterfeeding) Examples*. From <http://www.unc.edu/~moreton/Materials/Chainshifts.pdf>
- Moreton, E. (2004b). Non-computable functions in Optimality Theory. In J. McCarthy (Ed.), *Optimality Theory in Phonology: A Reader* (pp. 141-164). Malden, MA: Wiley-Blackwell.
- Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, 25(1), 83-127.
- Moreton, E., & Pater, J. (2012a). Structure and substance in artificial-phonology learning. Part I: Structure. *Language and Linguistic Compass*, 6(11), 686-701.
- Moreton, E., & Pater, J. (2012b). Structure and substance in artificial-phonology learning. Part II: Substance. *Language and Linguistic Compass*, 6(11), 702-718.
- Moreton, E., & Pertsova, K. (2014). Pastry phonotactics: Is phonological learning special? *Proceedings of NELS*, 43(2), 1-12.
- Moreton, E., & Pertsova, K. (2015). *Implicit and Explicit Processes in Phonological Learning*. Paper Presented at MFM 23 (Manchester Phonology Meeting), May 2015.
- Moreton, E., & Smolensky, P. (2002). Typological consequences of local constraint conjunction. (L. Mikkelsen, & C. Potts, Eds.) *Proceedings of WCCFL*, 21, 306-319.
- Morrisette, M., & Gierut, J. (2008). Innovations in the treatment of chain shifts. In *Optimality Theory, Phonological Acquisition and Disorders* (pp. 205-220). London: Equinox.
- Mortensen, D. (2006). *Logical and Substantive Scales in Phonology*. PhD Dissertation: University of California, Berkeley.
- Moshinsky, J. (1974). *A Grammar of Southeastern Pomo*. Berkeley: University of California Press.
- Nagle, T. (2008). Double opacity in Standard Colloquial Bengali verbs, through the lens of OT-CC. *Indiana University Working Papers*, 6, 35-53.
- Nagle, T. (2013). *Are Chain Shifts Generalizable? Wug-Testing Vowel Alternations in Bangla Present-Tense Verbs*. University of Indiana. Bloomington, Indiana: Unpublished Manuscript.
- Nagle, T. (2016). *The Implementation of Phonological Opacity in the Bangla Vowel Chain Shift*. Presented at FASAL 6, Amherst MA.
- Neasom, N. (2011). *On Constraining Phonological Terminology: The Limits of Synchronic Chain Shift*. MA Dissertation: UCL.

- Neasom, N. (2013). What is a synchronic chain shift? (D. Mazzarrella, I. Needham-Didsbury, & K. Tang, Eds.) *UCL Working Papers in Linguistics*, 25, 251 - 270.
- Ní Chiosáin, M. (1991). *Topics in the Phonology of Irish*. PhD Dissertation: University of Massachusetts, Amherst.
- Nibert, H. (1998). Processes of vowel harmony in the Servigliano dialect of Italian: A comparison of two non-linear proposals for the representation of vowel height. *Probus*, 10, 67-101.
- Niederer, B. (1998). *Les Langues Hmong-Mjen (Miáo-Yáo)*. Munich: Lincom Europa.
- Noske, R. (2012). The Grimm-Verner push chain in Contrast Preservation theory. In R. Noske, & B. Botma (Eds.), *Phonological Explorations: Empirical, Theoretical and Diachronic Issues* (pp. 63-86). Berlin: de Gruyter.
- Noske, R., Schinkel, J., & Smith, N. (1982). The question of rule-ordering. *Journal of Linguistics*, 18, 389-408.
- Oftedal, M. (1985). *Lenition in Celtic and in Insular Spanish: The Secondary Voicing of Stops in Gran Canaria*. Oslo: Universitetsforlaget.
- Orešnik, J. (1972). On the epenthesis rule in Modern Icelandic. *Arkiv for Nordisk Filologi*, 87, 1-32.
- Padgett, J. (2002). *Constraint Conjunction versus Grounded Constraint Hierarchies in Optimality Theory*. Retrieved from <http://people.ucsc.edu/~padgett/locker/conjunct.pdf>
- Parkinson, F. (1996). *The Representation of Vowel Height in Phonology*. PhD Dissertation: Ohio State University.
- Paster, M. (2004). Vowel height harmony and backing in Buchan Scots. *Phonology*, 21(3), 359-407.
- Pater, J. (1999). Austronesian nasal substitution and other \*NC Effects. In R. Kager, H. van der Hulst, & W. Zonneveld (Eds.), *The Prosody Morphology Interface* (pp. 310-343). Cambridge: Cambridge University Press.
- Pater, J. (2016). Universal grammar with weighted constraints. In J. McCarthy, & J. Pater, *Harmonic Grammar and Harmonic Serialism*. London: Equinox.
- Paul, H. (1896). Uber die Aufgaben der Wortbildungslehre. *Stizungsberichte der Konigl. bayer*, 692-713.
- Pierrehumbert, J. (2006). The statistical basis of an unnatural alternation. In L. Goldstein, D. Whalen, & C. Best (Eds.), *Laboratory Phonology VIII, Varieties of Phonological Competence* (pp. 81-107). Berlin: Mouton de Gruyter.
- Pierrehumbert, J., & Beckman, M. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Pitkin, H. (1984). *Wintu Grammar*. Berkeley: University of California Press.
- Poser, W. (1992). Blocking of phrasal constructions by lexical items. In I. Sag, & A. Szabolcsi (Eds.), *Lexical Matters* (pp. 111-130). Stanford: Center For the Study of Language and Information.
- Postal, P. (1969). Mohawk vowel doubling. *International Journal of American Linguistics*, 35(4), 291-298.
- Press, M. (1979). *Chemehuevi: A Grammar and Lexicon*. Berkeley: University of California Press.
- Prince, A., & Smolensky, P. (1993/2008). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: John Wiley & Sons.

- Puente, F. J. (1996). *A Bipolar Model of Height in Vowels and Consonants*. PhD Dissertation: University of Texas at Austin.
- Pullman, T. (2004). *The Morphophonology of Consonant Mutations in Irish*. MA Dissertation: Cambridge University.
- Pycha, A., Nowak, P., Shin, E., & Shosted, R. (2003). Phonological rule-learning and its implications for a theory of vowel harmony. (G. Garding, & M. Tsujimura, Eds.) *Proceedings of WCCFL*, 22, 423-435.
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Retrieved from R Foundation for Statistical Computing, Vienna, Austria: <http://www.r-project.org>
- Rawlins, K. (2006). Stress and epenthesis in Mohawk. Ms: USC Santa Cruz.
- Reber, A. (1967). Implicit learning of Artificial Grammar. *Journal of Verbal Learning and Verbal Behaviour*, 6(6), 855-863.
- Redington, M., & Chater, N. (1996). Transfer in Artificial Grammar Learning: A reevaluation. *Journal of Experimental Psychology: General*, 125, 123-138.
- Reid, T. (2010). *Aspects of Phonetics, Phonology and Morphophonology of Thok Reel*. Msc Dissertation: University of Edinburgh.
- Rice, K. (1993). A Reexamination of the feature [sonorant]: The status of 'sonorant obstruents'. *Language*, 69, 308-344.
- Rohlf, G. (1956-61) *Vocabolario dei Dialetti Salentini (Terra d'Otranto)* (3 vol) Munich: Beck'schen Verlagsbuchhandlung
- Rosenthal, R. (1979). The file-drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638.
- Round, E. (2010). *Kayardild Morphology, Phonology, and Morphosyntax*. PhD Dissertation: Yale University.
- Rubach, J. (1984). *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris Publications.
- Saffran, J., & Thiessen, E. (2003). Pattern induction by infant learners. *Developmental Psychology*, 39, 484-494.
- Sanders, N. (2003). *Opacity and Sound Change in the Polish Lexicon*. PhD Dissertation: UCSC.
- Savoia, L. (2015). *I Dialetti Italiani: Sistemi e Processi Fonologici Nelle Varietà di Area Italiana e Romancia*. Pisa: Pacini.
- Schmidt, D. S. (1996). Vowel raising in Basaa: A synchronic chain shift. *Linguistics*, 13, 239-267.
- Scott, N. (1957). Notes on the pronunciation of Sea Dayak. *Bulletin of the School of Oriental and African Studies*, 20, 509-512.
- Scott, N. C. (1964). Nasal consonants in Land Dayak. In D. Abercrombie (Ed.), *In Honour of Daniel Jones*. London: Longman.
- Shepard, R., Hovland, C., & Jenkins, H. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, 75(13), 1-42.
- Singh, R. (1987). Well-formedness conditions and phonological theory. In W. Dressler, H. Luschutzky, P. O., & J. Rennison (Eds.), *Phonologica 1984* (pp. 273-285). Cambridge: Cambridge University Press.
- Sluyters, W. (1988) Vowel Harmony, Rule Formats, and Underspecification: The Dialect of Francavilla Fontana in H. van der Hulst and N. Smith (Eds.) *Features, Segmental Structure, and Harmony Processes* (Part II) Foris: Dordrecht pp.161-184

- Smit, A. (1993). Phonologic error distributions in the Iowa-Nebraska articulation norms project: consonant singletons. *Journal of Speech and Hearing Research*, 36, 533-547.
- Smith, N. (1973). *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.
- Smith, N. (2010). *Acquiring Phonology: A Cross-Generational Case Study*. Cambridge: Cambridge University Press.
- Smolensky, P. (1993). *Harmony, Markedness, and Phonological Activity*. Paper Presented at the Rutgers Optimality Workshop 1, New Brunswick, NJ.
- Snow, K. (1963). A detailed analysis of articulation responses of 'normal' first grade children. *Journal of Speech and Hearing Research*, 37, 831-840.
- Spencer, A. (1999). *Chuckchee Homepage*. Retrieved from [http://privatewww.essex.ac.uk/~spena/Chukchee/CHUKCHEE\\_HOMEPAGE.html](http://privatewww.essex.ac.uk/~spena/Chukchee/CHUKCHEE_HOMEPAGE.html)
- Stockwell, R. (2002). How much shifting actually occurred in the Historical English Vowel Shift? In D. Minkova, & R. Stockwell (Eds.), *Studies in the History of the English Language: A Millennial Perspective* (pp. 267-281). Berlin: Mouton de Gruyter.
- Stockwell, R., & Minkova, D. (1988). The English Vowel Shift: Problems of coherence and explanation. In D. Kastovsky, & G. Bauer (Eds.), *Luick Revisited*. Tübingen: Gunter Narr.
- Swingle, K. (1993). The Irish and other mutations. (J. Mead, Ed.) *Proceedings of WCCDL*, 11, 451-466.
- Tessier, A.-M. (2012). Testing for OO-faithfulness in the acquisition of consonant clusters. *Language Acquisition*, 19(2), 144-173.
- Thomas, E. (2000). Spectral differences in /ai/ offsets conditioned by the voicing of the following consonant. *Journal of Phonetics*, 28, 1-25.
- Thomas, G. (2008). An analysis of the Xiamen tone circle. (N. Abner, & J. Bishop, Eds.) *Proceedings of WCCFL*, 27, 422-430.
- Thráinsson, H. (1994). Icelandic. In E. König, & J. van der Auwera (Eds.), *The Germanic Languages* (pp. 142-189). London: Routledge.
- Topolińska, Z. (1974). *A Historical Phonology of the Kashubian Dialects of Polish*. Berlin: Mouton de Gruyter.
- Torres-Tamarit, F., Linke, K., & van Oostendorp, M. (2016). *Approaches to Metaphony in the Languages of Italy*. Berlin: Mouton de Gruyter.
- Trommer, J. (2009). Chain-Shifting Mutation in Irish and Multi-Valued Features. *OCP 6, Edinburgh*.
- Trommer, J. (2011). *Phonological Aspects of Western Nilotic Mutation Morphology*. Habilitation thesis: University of Leipzig.
- Trudgill, P., Gordon, E., & Lewis, G. (1998). New dialect formation and Southern Hemisphere English: The New Zealand short front vowels. *Journal of Sociolinguistics*, 2, 35-51.
- Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition: Methods, Theory, and Practice* (pp. 141-178). Washington, DC: Georgetown University Press.
- Ullman, M., Corkin, S., Coppola, G., Hickok, G., Growdon, J., Koroshetz, W., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules

- are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9(2), 266-276.
- Urbańczyk, S. (1972). *Zarys Dialektologii Polskiej*. Warsaw: PWN.
- Valente, V. (1975). *Puglia*. Pisa: Pacini.
- van Driem, G. (1987). *A Grammar of Limbu*. Amsterdam: de Gruyter.
- van Oostendorp, M. (2004, February 11). *Phonological Recoverability in Dialects of Dutch*. Retrieved October 20, 2013, from <http://www.vanoostendorp.nl/pdf/recoverable.pdf>
- van Oostendorp, M. (2005). Expressing inflection tonally. *Catalan Journal of Linguistics*, 4(1), 107-127.
- Vaux, B. (2008). Why the phonological component must be serial and rule-based. In A. Nevins, & B. Vaux, *Rules, Constraints, and Phonological Phenomena*. Oxford: Oxford University Press.
- Velleman, S. (1988). The role of linguistic perception in later phonological development. *Applied Psycholinguistics*, 9, 221-236.
- Walker, R. (2005). Weak triggers in vowel harmony. *Natural Language and Linguistic Theory*, 23, 917-989.
- Walton, F. (2012). *Chain Shifts in First Language Acquisition*. MA Dissertation: Utrecht University.
- Wang, F. (1985). *Miao yu Jianzhi [Sketch Grammar of the Miao Language]*. Beijing: Minzu Chubanshe.
- Wang, W. (1967). Phonological features of tone. *International Journal of American Linguistics*, 33(2), 93-105.
- Wells, J. (1982). *Accents of English (Three Volumes)*. Cambridge: Cambridge University Press.
- Wheeler, M. (2005). *The Phonology of Catalan*. Oxford: Oxford University Press.
- White, J. (2013). *Bias in Phonological Learning: Evidence from Saltation*. PhD Dissertation: UCLA.
- White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, 130(1), 96-115.
- White, J., & Sundara, M. (2014). Biased generalization of newly learned phonological alternations by 12-month-old-infants. *Cognition*, 133(1), 85-90.
- Williams, J. (2005). Learning without awareness. *Studies in Second Language Acquisition*, 27, 269-304.
- Williamson, K. (1977). Multivalued features for consonants. *Language*, 53, 843-871.
- Wilson, C. (2003). Experimental investigations of phonological naturalness. (G. Garding, & M. Tsujimura, Eds.) *Proceedings of WCCFL*, 22, 101-114.
- Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, 30(5), 945-982.
- Wolf, M. (2007a). What constraint connectives should be permitted in OT? *University of Massachusetts Occasional Papers in Linguistics*, 36, 151-179.
- Wolf, M. (2007b). For an autosegmental theory of mutation. *University of Massachusetts Occasional Papers in Linguistics*, 32, 315-404.
- Wolf, M. (2011). Limits on global rules in Optimality Theory with Candidate Chains. *Phonology*, 28, 87-128.
- Zager, D. (1980). *A Real-Time Process Model of Morphological Change*. PhD Dissertation: Buffalo SUNY.

- Zhang, J., Lai, Y., & Sailor, C. (2006). Wug testing the "tone circle" in Taiwanese. (D. Baumer, D. Montero, & M. Scanlon, Eds.) *Proceedings of WCCFL*, 25, 453-461.
- Zimmermann, E. (2014). *A Phonological Account of Morphological Length*. PhD Dissertation: Universität Leipzig.
- Zuraw, K. (2002). Vowel reduction in Palauan reduplicants. *MIT Working Papers in Linguistics* 44, 8, p.  
<http://www.linguistics.ucla.edu/people/zuraw/dnldpprs/PalRedup.pdf>.