

UC Berkeley
Dissertations, Department of Linguistics

Title

Logical and Substantive Scales in Phonology

Permalink

<https://escholarship.org/uc/item/0sp1b9w8>

Author

Mortensen, David

Publication Date

2006

Logical and Substantive Scales in Phonology

by

David Roland Mortensen

B.A (Utah State University) 2000

M.A. (University of California, Berkeley) 2003

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in

Linguistics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Sharon Inkelas (co-chair)

Professor James A. Matisoff (co-chair)

Professor Larry M. Hyman

Professor Johanna Nichols

Fall 2006

Abstract

This study addresses two assumptions that have accompanied generative phonology from its inception, namely the binarity and phonetic grounding of features. It argues, based on both theoretical and empirical grounds, for the existence of n -ary phonological relationships that are not necessarily grounded in a phonetic dimension or parameter. It presents, further, a model labelled Structural Optimality, which combines representations of this type with Optimality-Theoretic grammars.

Four types of phenomena are presented as arguments for Structural Optimality: chain shifts, circle shifts (i.e. circular chain shifts), phonologically-driven ordering effects in coordinate compounds, and certain dissimilatory effects in reduplication. Each of these phenomena present problems for theories that are inseparably tied to binarity, strongly committed to phonetic grounding, or both.

Chain shifts have conventionally been hard to capture as a single process and have presented difficulties to output-oriented phonological models like Optimality Theory. This study raises another difficulty: phonetically grounded chain shifts may become more and more arbitrary over time without losing their original structure. Structural Optimality presents a solution to both of these problems by treating chain shifts as traversals of scales, the essential component of which is their logical structure rather than their phonetic manifestation. This is done by invoking a novel concept: directional anti-identity constraints, the class of constraints that require that elements be higher on a scale than the elements to which they correspond.

An even more difficult problem for earlier theories of phonology has been that of circle shifts. For Optimality Theory, especially, these have been challenging since one of the demonstrable properties of classical OT, *Harmonic Ascent*, “all unfaithful mappings must be markedness-reducing,” explicitly rules out such patterns. This study demonstrates that circle shifts are actually more widespread than has been traditionally believed, are as phonological in nature as other phenomena addressed by phonologists, and therefore must be engaged by phonological theory. Further, it demonstrates that the predictions made by contrast-preservation theories of circle shifts—particularly, that all circle shifts must involve neutralization—are not correct, but that directional anti-identity, as motivated by ordinary chain shifts, actually predicts the existence of circle shifts as well. That circle shifts may not involve neutralization is demonstrated by surveying the diverse set of tone circles found in dialects of Southern Min (a Chinese language). Structural Optimality analyses of circle shifts from Xiamen Chinese, A-Hmao, and Jingpho are presented, showing how this model can capture all of these—very different—patterns. The case of Jingpho is also seen to motivate general scalar anti-identity. Arguing for these many cases, this study asserts that Harmonic Ascent is not a property of natural languages.

Directional anti-identity is also shown to predict the existence of another phenomenon, when applied to correspondence relationships across a string (rather than between input and output): the order of coordinate compounds (co-compounds), being free from syntactic requirements on sequence, can be ordered according to n -ary phonological criteria of the same type that produce chain shifts and circle shifts. Numerous examples of effects of this kind are presented, as are analyses couched in Structural Optimality. The case of Hmong (Mong Leng) co-compounds is treated in particular detail, demonstrating that the ordering phenomenon in that language simultaneously argues for optimality grammars and against phonetic grounding or substantial optimization.

This part of the study further demonstrates how morphotactic patterns can be brought to bear in the study of phonological grammars and representations, and it demonstrates the striking parallel between two apparently unrelated phenomena: chain-shifting and phonologically-driven co-compound ordering.

Finally, this study ties yet another phenomenon to the rest: echo reduplication. It shows how the same sets of constraints that are motivated by circle-shifting and co-compound ordering play a role in the “dissimilatory” effects in echo-reduplication. This is shown by looking at examples from English, Jingpho, and Eastern A-Hmao. The A-Hmao case is particularly enlightening, showing a three-way patterning in dissimilation that is easy to capture in terms of scales but more difficult to capture using traditional features.

The upshot of the study is that a descriptively and explanatorily adequate theory of phonology must include representations that are neither strictly binary nor strictly grounded in phonetic parameters. It provides empirical evidence for the kinds of relationships posited by early structuralists on theory-internal grounds. In doing so, it frees Universal Grammar from the shifting moorings of substance and allows the theory of phonology to be a theory of grammar rather than a theory of motor control or psychophysics.

for Rahel

Light to my shadows,
dawn to my night.

Contents

Contents	ii
Acknowledgements	viii
1 Introduction	1
1.1 Precis	3
1.1.1 Survey of previous work	3
1.1.1.1 Phonological relationships as logical abstractions	3
1.1.1.2 Scalar representations	5
1.1.2 Structural Optimality: an overview	7
1.1.2.1 Optimality Theory	8
1.1.2.2 Logical scales	10
1.1.2.3 Scale-referring constraints	11
1.1.3 Empirical motivations	15
1.1.3.1 Chain shifts and circle shifts	15
1.1.3.2 Ordering effects	17
1.1.3.3 Graded dissimilation and other scalar effects in reduplication	18
1.2 Overview of Chapters	19
2 Logical Scales and the Theory of Structural Optimality	21
2.1 Structural Optimality	21

2.2	The Definition and Structure of Logical Scales	22
2.2.1	Natural and Unnatural Scales	28
2.2.2	Logical Scales and Richness of the Base	29
2.3	Structural Harmony Constraints	31
2.3.1	ENDMOST	32
2.3.2	TOP	34
2.3.3	EXTREME	34
2.4	String Internal Correspondence	35
2.4.1	Properties of correspondence relationships	39
2.5	Syntagmatic Relational Harmony Constraints	40
2.5.1	PLATEAU	42
2.5.2	NOPLATEAU	43
2.5.3	WAX and NOWAX	45
2.6	Input-Output Relational Harmony Constraints	48
2.6.1	SAME	48
2.6.2	DIFF	49
2.6.3	HIGHER and NOHIGHER	53
2.7	A Factorial Typology of Mapping Schemas	55
2.7.1	The Identity Mapping	56
2.7.2	Neutralization	56
2.7.3	Neutralization with “Bounce-Back”	57
2.7.4	Chain shift	60
2.7.5	Ring (circular chain shift)	62
2.7.6	Comparison of Typological Predictions	66
3	Chain Shifts and Scales	68
3.1	Chain Shifts: Problems and Prospects	68

3.2	Western Hmongic Tonal Chain Shifts	71
3.2.1	Dananshan Hmong	72
3.2.1.1	Origin of the Chain Shift	74
3.2.1.2	Problems with an Autosegmental Account	75
3.2.2	Eastern A-Hmao	82
3.2.3	Mashan Dialects	84
3.2.3.1	Naturalness, productivity, and the proper scope of phonology	90
3.3	Fricative Place Chain Shift in Acquisition	93
3.4	Conclusion	96
4	Circular Chain Shifts	98
4.1	Southern Min Tone Sandhi Circles	100
4.1.1	Xiamen and Mainstream Taiwanese	101
4.1.1.1	A digression on “psychological reality”	103
4.1.1.2	Allomorph-selection hypotheses	106
4.1.1.3	Rule based analyses	108
4.1.1.4	Contrast preservation analyses	110
4.1.2	Checked-syllable (rusheng) circles	115
4.1.2.1	Dong-an	115
4.1.2.2	Tainan and other Taiwanese dialects	117
4.1.2.3	Chaoyang	118
4.1.2.4	Chao-an	118
4.1.3	Chaozhou (Chaoshan)	120
4.1.4	Yilan: a circle shift sans neutralization	121
4.1.5	Dongshan	122
4.1.5.1	A Structural Optimality account	123

4.1.5.2	Relevance for contrast-preservation theories	129
4.1.6	The historical evolution of circle shifts in Southern Min	130
4.1.6.1	First principles	130
4.1.6.2	Xiamen and Mainstream Taiwanese	133
4.1.6.3	Dongshan	135
4.2	A-Hmao Bounce-back Effects	136
4.2.1	Western A-Hmao	137
4.2.2	Eastern A-Hmao	144
4.3	Jingpho Bounce-back Effect	152
4.3.1	With initial checked syllables	153
4.3.2	With initial smooth syllables	156
4.4	Circle Shifts and Bounce-back effects as counter-examples to harmonic ascent	168
5	Ordering Effects in Coordinate Compounds	174
5.1	Coordinate Compounds	175
5.1.1	Co-compounds in Mainland Southeast Asia	177
5.1.2	The relation of co-compounds to other constructions	185
5.1.2.1	Copulative compounds	185
5.1.2.2	Irreversible binomials	187
5.1.2.3	Pseudo-reduplicative compounds and echo-reduplication constructions	191
5.2	Survey of Co-compound ordering effects	205
5.2.1	Soft (stylistic) ordering effects	205
5.2.2	Hard (grammatical) ordering effects	208
5.3	Co-Compounds in Chinese	210
5.4	Co-Compounds in Lahu	215

5.5	Co-Compounds in Jingpho	221
5.5.1	Ordering effects	222
5.5.1.1	Data	223
5.5.1.2	Analysis	224
5.5.2	Comparison to Tangkhul	230
5.6	Co-Compounds in Hmong (Mong Leng)	235
5.6.1	The ordering generalization	237
5.6.2	Factoring the observed scale	245
5.6.2.1	Independent evidence for the “register” scale	245
5.6.2.2	A competition between two WAX constraints	246
5.6.3	Explaining the origin of the scales	255
5.7	Co-Compounds in Qe-Nao	258
5.8	Co-Compounds in Qo-Xiong	261
5.9	The Significant of Coordinate Compounds for Phonology	268
6	Logical Scales and Echo Reduplication	271
6.1	English	272
6.1.1	Echo-reduplication and Morphological Doubling Theory	272
6.1.2	A preliminary analysis of English echo-reduplication	275
6.2	Jingpho	278
6.2.1	Jingpho echo-reduplication and co-compounding	278
6.2.2	Explaining Jingpho echo-reduplication	280
6.3	Eastern A-Hmao	284
6.3.1	Phonological patterning in Eastern A-Hmao reduplication	284
6.3.2	A-Hmao reduplication as a scalar phenomenon	286
6.3.3	An analysis of A-Hmao echo reduplication	288
6.4	Conclusion	291

7 Conclusion and Directions for Further Research	292
7.1 Conclusions	292
7.1.1 Empirical phenomena and analyses	293
7.1.2 General theoretical implications	296
7.2 Remaining Questions and Directions for Further Research	298
7.2.1 Bias toward Southeast Asian tone	298
7.2.2 Logical scales and other features	299
7.2.3 Formal properties of Structural Optimality	300
7.2.4 Scales outside phonology	301
Bibliography	302

Acknowledgements

Acknowledging the contributions of others is far from the easiest task involved in writing a dissertation. If one is too generous, thanking everyone possible, one runs the risk of trivializing these expressions of gratitude; if one is not generous enough, one might fail to credit individuals who have made vital contributions to the work. I hope that my attempt to navigate this strait will neither be over-thorough nor offensively stingy.

My time as a graduate student, and particularly the period in which this dissertation has been completed, has been for me a time of unprecedented happiness and excitement but—even more so—a time of great pain and despair. Credit for helping me complete this dissertation in its current form goes as much to those who held me back from the gaping abyss of oblivion as to my intellectual and academic guides.

This must be mentioned because my thesis advisers—Sharon Inkelas and James (“Jim”) Matisoff—have supported me in both ways. They have not only inspired, interrogated, and helped to shape most of the ideas and analyses I present here; they have also—each in their own way—helped me live through some very dark times, showing to water what is due only to blood. Jim inspired my interest in many of the languages and phenomena that are discussed in this study and his ability to make uncanny connections between ideas and facts has informed and improved my work in ways too numerous to count. However, it is for lunches and Blake’s on Telegraph when life seemed unbearable that I will remember him most. As for Sharon, the theoretical ideas around which this dissertation is built would have never been developed if she had not listened patiently as I sat in her office and stuttered on about “abstract scales” late one

morning in the spring of 2003. I must thank her for the many hours of conversation and endless words of encouragement without which it is doubtful that I would have written this dissertation. More than any of that, though, I am grateful to her for allowing me to be an honorary member of her family through a difficult summer in Boston during 2005.

I am also deeply indebted to the rest of my committee. Larry Hyman, in fact though not in name, has served as yet a third adviser to me, not just with regard to my dissertation, but in my academic life as a whole, a gift for which I will be endlessly grateful. I must thank Larry for keeping his office door open, and for being ready and willing to talk about tone at a moment's notice (at any moment of night or day). I must express my thanks to Johanna Nichols too, who radiated encouragement that helped me push through the last stages of my writing and helped me see my research from a broader, more meaningful, perspective.

Other members of the Berkeley faculty provided essential feedback on presentations of my dissertation research that have improved the quality of this study. Prominent among these are Keith Johnson, Ian Maddieson, and Line Mikkelsen. Among my student colleagues, I must also thank Teresa McFarland, Anne Pycha, Christian DiCanio, Nicholas Fleisher, Jisup Hong, Yuni Kim, Mary Paster, Gabriela Caballero, Angela Strom-Weber, Maziar Toosarvandani, and Dominic Yu for many thought-provoking discussions and warm, personal friendship. I would especially like to recognize Kenneth VanBik, my fellow Tibeto-Burmanist and STEDTnik, for friendship and scholarly exchange starting from my first day at Berkeley.

Beyond the Berkeley fold, I must thank Juliette Blevins for many interesting conversations which had a great influence on how I understand phonology. Also, my gratitude goes to Alan Prince: it was two brief but stimulating conversations with him that led me to change the topic of my dissertation in the middle of the writing process. While

this was doubtless not Alan's intent, he helped me to write a much better and more interesting dissertation than I would have otherwise.

Credit for this work must also go to those individuals who helped me maintain some semblance of sanity through the most painful months of my life. My parents, Gene and Shirley Mortensen, rank high among these, as do my many siblings. Their patient support leaves no doubt in my mind regarding their love and loyalty. Further, I must mention Sun Kim for late night trips to Nation's when he found I had not eaten; Reed Asay, for always being around to cheer me up; Sam Petty and Beth Heath, for keeping watch with me through difficult nights; Dan Volmar, for sympathetic commiseration, Jeff Pynes, for reminding me that life and linguistics were still exciting; Dominic Yu, for repeatedly accompanying me to Naan'N'Curry even though the food was too spicy for him; Dean Criddle, for wise counsel and a listening ear; and Glenn Worden, for being a friend without peer (and for proofreading the entire dissertation). Belen Flores, too, was always there to shepherd me through the University bureaucracy, and was always an understanding friend. She bailed me out of scrape after scrape (and in one last bail-out, will be the one to file this dissertation with the UCB Graduate Division). I extend my heartfelt thanks to all of you.

Finally, I would also like to give thanks to the shimmering and inscrutable hand of Divinity, which has caressed my face in my hours of deepest despair, and has given me the ultimate comfort in the final months of my dissertation-writing—the love of Rahel Ringger, to whom this work is dedicated.

Less personally, portions of the research on which this work is based were supported by grant BCS-9904950 from the National Science foundation and grant PA-24168-02 from the National Endowment for the Humanities.

Chapter 1

Introduction

Me thynketh it acordaunt to resoun
To tell yow al the condicion
Of ech of hem, so as it seemed to me,
And whiche they weren and of what
degree.

Canterbury Tales, "Prologue"
Geoffrey Chaucer

This study attempts to advance two main ideas: that relationships between the sounds in a language may be scalar rather than simply binary, and that these relationships may not be strongly grounded in a phonetic dimension. These ideas strike at the heart of much of the work in phonological theory done since the advent of generative phonology. On the other hand, these notions also arise as a natural consequence of the endeavor embraced by modern phonological theory. Neither of these ideas is wholly novel, but the combination of these two ideas in the manner presented in this work is. I show here that this combination provides a new perspective on a number of phenomena in phonology that have not been adequately studied in the past, quite possibly because the theoretical framework for understanding these patterns did not exist in a fully elaborated form.

The theoretical framework developed in this work is called Structural Optimality. It consists of a system of constraints over logically-defined phonological scales which operate within the context of an Optimality Theory grammar. Specifically, it requires *scales* which group all phonological entities of some type (for example, all vowels)

into a (potentially language-specific) hierarchy and *constraints* that evaluate relationships between entities belonging to these scales. One set of constraints apply only to outputs and require that elements in some scale be at a particular position in that scale. Another set evaluates relationships between corresponding elements belonging to a scale, with constraints requiring that elements be at the same point on a scale as their correspondents, at a different point on the scale as their correspondents, at a higher point on the scale than their correspondents, and at no higher a point on the scale than their correspondents. The theory thus includes not only a concept of identity (really, equivalence) but also a concept of anti-identity and the novel idea of “directional anti-identity.” An examination of empirical phenomena shows that it is necessary to refer to all of these types of relationships in a descriptively and explanatorily adequate theory of phonological grammar.

Phenomena analyzed in this study include chain shifts, circular chain shifts, coordinate compound ordering effects, and various phonological alternations in reduplication. Certain of these phenomena have been difficult (chain shifts, graded-dissimilation in reduplication) or impossible (circular chain shifts) to capture in conventional versions of Optimality Theory, and even other, less restrictive, models of phonological theory. The intent of this study, in these areas, is not only to show how Structural Optimality accounts for these patterns, but to provide a general treatment of these poorly-understood phenomena.

In §1.1 I will provide a condensed overview of the intellectual background, theoretical formalism, and empirical coverage of Structural Optimality. Then, in §1.2, I will provide a chapter-by-chapter outline of this study.

1.1 Precis

Even the most radical sounding ideas presented in this study are adapted from concepts with a long historical in linguistic science. The earlier work upon which Structural Optimality builds is briefly surveyed in §1.1.1. Following that review of relevant literature, in §1.1.2, I will present a brief and simple overview of the Structural Optimality formalism, intended to prepare the reader for the more detailed and technical introduction to the framework given in Chapter 2. Finally, in §1.1.3, I will introduce the empirical phenomena that Structural Optimality is meant to explain. These same phenomena are treated more fully in Chapters 3 through 6. This preliminary presentation is meant to introduce the reader to the fundamental extra-theoretical issues addressed in this study.

1.1.1 Survey of previous work

While some of the ideas in this study may seem unorthodox, all of them are grounded in ideas with a long history in linguistic and phonological theory. It is appropriate, before introducing the theory in all of its details, to provide a brief account of the history of ideas that lead to the development of Structural Optimality and to acknowledge the work of earlier scholars in the same vein of research.

1.1.1.1 Phonological relationships as logical abstractions

Modern linguistic theory began with structuralism, and structuralism—in many ways—started with the set of ideas that gave phonology the phoneme (Baudouin de Courtenay (1895 [1972])). From the time of de Saussure (1916), and apparently before, language was conceived as a structure of units that were defined by their relationship to one another. Linguistic signs, it was claimed, consisted of arbitrary relationships between signifiers and signifieds; phonological relationships were similar in that they were understood in terms of systems of contrasting elements rather than in terms of their sub-

stance. This view of phonology reached its height in the work of Hjelmslev (1961), though it was also present in American structuralism as attested by work such as Harris (1951). In more modern form, this structuralist view also served as an underpinning for work like Jakobson et al. (1951). Furthermore, a radically abstract position like that of Hjelmslev has been advanced by certain contemporary linguists coming out of the generative tradition, particularly Hale and Reiss (2000).

However, since Chomsky and Halle (1968), generative phonologists came to adopt psychologized version of a Prague-school view of phonology in which phonological features were more than just classificatory devices, but rather were in universal relationships with phonetic substance, as expressed by Roman Jakobson (Jakobson and Waugh 1979). Generative phonologists like Postal (1968) aggressively advocated this point of view as superior to the traditional structuralist understanding of phonological relationships.

This view of phonological features as perceptual primitives or as instructions to articulators prepared the way for the entry of phonetics into theoretical phonology on a much larger scale. Although SPE and other early generative work assigned only a small role to phonetic substance in the phonology—defining the set of phonological features—this precedent opened the doors to a reversal of the structuralist tradition and a return of phonetics into phonology. This reversal was protested by linguists like Foley (1970, 1972, 1977) but it seems to have paved the way for an even more radical incursion of phonetics into the phonological domain with the emergence of theories of *natural phonology*, that is, Natural Phonology and Natural Generative Phonology (Stampe 1973; Hooper 1976). Again, despite rejoinders like Anderson (1981), this movement advanced by fits and starts to become one of the most important trends in theoretical phonology at the time of writing. Archangeli and Pulleyblank (1994) argued very persuasively for the phonetic grounding of phonology, and Optimality Theory (Prince

and Smolensky 1993) has—from its inception—been associated with the claim that substantive optimization (and therefore phonetics) drives phonological processes. This position has been expressed in radical form by a considerable number of influential phonologists (Flemming 1995, 2004; Steriade 1997, 2002; Hayes and Steriade 2004; Boersma 1998; Hayes 1996; Hayes and Steriade 2004). Such theories undermine the distinction between phonetics and phonology, and, in some cases, they deny its existence entirely.

However, this position has not been without its critics. On the one hand, there is a considerable body of experimental work demonstrating that language users show little or no preference for “natural” phonological processes—those that are grounded in phonetic principles and share properties with cross-linguistically common phonological patterns—over processes that are similarly complex but of a phonetically arbitrary nature (Pycha et al. 2003; Seidl and Buckley 2005)¹. An even larger body of work argues that typological patterns in phonological processes are the result of facts about the historical development of phonologies, rather than the direct phonetic grounding of phonological grammars (Ohala 1974, 1981, 1993, 1995b; Hyman 2000; Blevins and Garrett 1998, 2004; Blevins 2004; Kavitskaya 2002; Barnes 2002; Mielke 2003, 2004b,a). While not always viewed this way by those working in this program, such work opens the door for the reemergence of a pure structural phonology. The current work builds on that foundation.

1.1.1.2 Scalar representations

Within generative phonology, it has been conventional to treat all phonological features as binary or (more recently) privative. In both cases, surface oppositions in phonology are binary in nature—plus or minus feature values, presence or absence of a privative

¹But see also Wilson (2003).

feature. The rationale for binary features was given by Chomsky and Halle (1968) and was almost universally accepted by later generative phonologists. However acceptance was not total and there is a long counter-tradition of arguing for multi-valued features. Among the early objectors to strict binarity were Contreras (1969), Rivas (1977), Stahlke (1977), Foley (1977), Williamson (1977), and Lindau (1978). Of these proposals, that of Foley bears the closest resemblance to the system presented in this study, combining both phonological abstraction and scalar representations.

More recent work has also argued for scales in phonology. Prince and Smolensky (1993) present a way of encoding a sonority scale in grammatical constraints. This view of sonority has a long history dating back to Sievers (1881) and has been continued by scholars such as Jespersen (1904), Vennemann (1972), Foley (1977), and Selkirk (1984). Optimality Theory, too, with its system of ranked and violable constraints, was a natural context for incorporating phonological scales into the grammar. This line of theorizing was further pursued in the work of de Lacy (2002a,b).

However, prior to the appearance of de Lacy's work, other scalar proposals that were couched in Optimality Theory appeared. The most developed of these was Gnanadesikan (1997), which employed scalar features directly and allowed constraints that both applied to individual points on a scale and constraints that regulated relative positions on the scale. In her system, all scales were actually ternary features, and were meant to replace corresponding binary features. Gnanadesikan applied her system to a number of different phenomena, but the most important of these was Celtic consonant mutation (*eclipsis*). Proposing a ternary "inherent voicing scale," she used her formalism to explain the chain shift from voiceless oral to voiced oral and from voiced oral to nasal.

Gnanadesikan's proposal was representationally oriented. While it did add constraints to the grammar, the largest innovation was revising the featural makeup of segments. In contrast, de Lacy (2002a,b) presented a view of phonological scales that

was almost completely non-representational, as least under one interpretation. In this theory, scales (which could be of arbitrary arity) acted more as meta data, defining which constraints could exist but not acting like normal features. Instead of being the primitives of which complex structures like segments are composed, these scales say something about the relationship between non-primitive elements like segments. A similar view of scales is taken up in this study, following de Lacy. Scales are not treated as featural primitives but as structures that relate feature geometries, segments, rimes, and so forth.

Unlike the system presented in this work, the scalar relationships employed in de Lacy's framework are expressed directly in the constraints. This encoding takes the form of a stringency hierarchy. Any constraint that penalizes a "less-marked" structure also penalizes any structure more marked. For example, there is a constraint $\delta_{Ft} \leq \{i,u\}$ which penalizes high vowels in head positions (because of their low sonority) if there is a similar constraint against mid vowels, this constraint would have to penalize mid vowels in this location as well: $\delta_{Ft} \leq \{i,u,e,o\}$. This implementation of phonological scales makes fixed rankings of these constraints unnecessary and allows for "conflation," where elements that are at different points on some universal scale because as if they were at the same point in the grammar of one particular language. Both of these properties are required, though, only because the scales in de Lacy's theory are meant to be universal. The scales in the current study, as will be seen, are held to be learner-constructed and language specific, making stringency and unnecessary complication.

1.1.2 Structural Optimality: an overview

Chapter 2 provides a detailed technical overview of logical scales, Structural Optimality, and the accompanying formalism. However, it is anticipated that this presentation will be too dense and technical to meet the needs of readers who are concerned primarily with understanding a particular analysis in the more empirically-oriented sections of

the study. For these readers, and for all readers who are more interested in the way that logical scales are applied than in their theoretical motivations and the technical details of their implementation, the following introduction is provided. It should be sufficient to allow the reader to follow most of the analyses given in Chapters 3 through 6, though it still may be necessary to refer to Chapter 2 for clarification of certain minor details.

As discussed above, Structural Optimality is an outgrowth of Optimality Theory (Prince and Smolensky 1993). There are many good introductions to this framework and it is not necessary to outline the full theory here, but I nevertheless will provide a very brief introduction in §1.1.2.1 in the hope that someone will read this work long after OT has been consigned to the dustbin of intellectual history. Situated within OT, Structural Optimality has two components, one representational and one operational. The representational component is a system of logical scales, which will be described in § 1.1.2.2. The operational component is a system of ranked, violable constraints that refer to these scales. These will be described in §1.1.2.2.

1.1.2.1 Optimality Theory

Optimality Theory (Prince and Smolensky 1993) or OT is a theory of grammar in which an optimal output is selected from a larger set of candidates by being the most harmonic candidate relative to a hierarchy of ranked, violable constraints. The grammar consists of three components which are called GEN, CON, and EVAL. The function of GEN is to produce all of the potential outputs for a particular input. CON is a hierarchy of ranked, violable constraints. The symbol \gg is used to represent domination—the state where one constraint outranks another. For any two constraints CONSTRAINT 1 and CONSTRAINT 2, wither CONSTRAINT 1 \gg CONSTRAINT 2 or CONSTRAINT 2 \gg CONSTRAINT 1. EVAL takes the candidates generated by GEN and chooses the best candidate relative to CON.

The selection of the optimal candidate proceed according to a very simple algo-

rithm: Look at the first constraint. Find the minimum number of violations of this constraint and eliminate all candidates that violate the constraint more times than this minimum. Iterate through the constraints in this manner, descending the hierarchy, until only one candidate remains. This is the optimal candidate and is, therefore, the chosen output.

This operation is depicted using a notational device called a tableau. The tableau depicts each of the candidates in rows and each of the constraints in columns. Constraint violations are marked with asterisks and the elimination of a candidate is indicated with an exclamation point. Winning candidates are marked with a pointing finger:

(1) **Example OT Tableau**

	CONSTRAINT 1	CONSTRAINT 2
(a) bad	*!	*
(b) better	*!	
☞ (c) best		*

Of course, a tableau cannot list all of the candidates that would be generated by GEN, so it is the responsibility of the OT practitioner to identify those candidates that are the strongest competitors with the winning candidate.

In classical Optimality Theory, constraints are of two types, conventionally called markedness and faithfulness constraints. Markedness constraints, which might be better called *structural harmony constraints* evaluate aspects of output candidates without making reference to inputs. Faithfulness constraints evaluate the relationships between inputs and outputs. They do this by penalizing differences between the input and an output candidate. This is implemented using a concept called correspondence. When the situation is examined at sufficient depth, candidates are not simply output strings. Instead, they are input-output pairs that include a correspondence relationship between the input and the output. This relation determines which segments, features, and so

forth in the output are “linked” to which counterparts in the input. Faithfulness constraints make reference to both input and output in conjunction with this correspondence relation.

There is much more that could be said about OT. The presentation given here is only meant to orient the reader towards some of the most important conceptual, notational, and terminological aspects of this framework.

1.1.2.2 Logical scales

The Optimality Theory formalism is a theory of grammars and not a theory of representations. The current study, and the theory of Structural Optimality, rest largely on the introduction of a representational device and the integration of this device into the grammatical machinery of OT. This representational device is the logical scale. Such constructs are scales because they have more than two points: they are like a multi-valued feature. They are logical in the sense that they encode a logical relationship between phonological constituents—features, feature structures, segments, and so forth—rather than encoding some continuum from the gross phonetic world.

One example of such a scale is sonority scale over the vowels:

(2) SONORITY SCALE

$$S = \{i,u\} < \{e,o\} < \{a\}$$

This means that /i/ and /u/ are at the same on the scale but are lower on the scale than /e/ and /o/, which are also at the same step as one another but are lower on the scale than /a/. Note that we will assume that the scale includes all of the possible vowels but that only the vowels that are of immediate relevance are listed. This is a (necessary) notational choice with no theoretical significance. The interesting aspect of these scales is that they need not mirror an easily recognizable phonetic dimension. For example, the following is also a possible scale:

(3) VOWEL SCALE

$$S' = \{e,o\} < \{i,u\} < \{a\}$$

In fact, any ordering of elements of a single type (for example, segments, consonants, vowels, tones, and so forth) is a possible scale in Structural Optimality. The scales are characterized by their logical structure, explored further in § 2.2, rather than in terms of the relationship between a position on the scale and some phonetic property.

By their very nature, logical scales must be seen as learner-constructed and language-specific. This mirrors the idea that has been advanced by various phonologists that many apparent phonological universals and relationships between phonetics and phonology are the result not of some property of an innate linguistic competence but arise instead from the facts of language learning and language change.

1.1.2.3 Scale-referring constraints

Scales, as a representational device, need some interface with the grammar. This is provided by scale-referring constraints (to borrow a term from de Lacy (2002a)). Just as there are two types of constraints in classical OT, there are two types of constraints in Structural Optimality. Those of the first type are *structural harmony constraints* that, like markedness constraints, simply evaluate output structures without making reference to any correspondence relation. Those of the second type are *relational harmony constraints*. These constraints, like faithfulness constraints, evaluate strings in conjunction with a correspondence relation. Relational harmony constraints, in turn, exist in two flavors. One evaluates an input-output pair in conjunction with a correspondence relation between the input and the output (like a conventional faithfulness constraint). The other evaluates only an output string in conjunction with a correspondence relation between the elements in it. This second type of correspondence has already been explored by investigators such as Walker (2000a,b); Hansson (2001); Rose and Walker

(2004)

The proposed structural harmony constraints may be defined as follows:

(4) a. ENDMOST[*S*]

For any element from the scale that is present in the output, there are no steps on the scale between that element and the low end of the scale.

b. TOP[*S*]

Any element from the scale that is present in the output is at the top (high end) of the scale.

c. EXTREME[*S*]

An element from the scale that is present in the output is not in the middle of the scale (that is, there are not steps on both sides of it).

The constraint ENDMOST acts as if it is “gradiently violable.” That is to say, it returns one violation for each step on the scale between the offending element and the low end of the scale. We can view it, however, as a constraint against such steps on the scale (that reside between some element in the scale and the low end of the scale) rather than against the elements themselves and, in so doing, conceive of the constraint in categorical rather than gradient terms. The other two constraints, TOP and EXTREME, are unambiguously categorical.

The relational constraints have the virtue of being derivable wholly and simply from aspects of the definition of the scale and negation. As will be discussed in greater detail below, a scale is a order over a set which is antisymmetric under equivalence. This means that if a is less than or equal to b and b is less than or equal to a , then a and b are equivalent. For such an ordering, there are two essential relations: the ordering relation, which is written as \leq , and the equivalence relationship that is written as \equiv . The relational constraints are statements about these relations:

(5) **Relational scale-referring constraints**

<i>Type</i>	<i>Relation</i>	<i>String-internal</i>	<i>Input-output</i>
<i>directional identity</i>	$a \leq b$	NOWAX	NOHIGHER
<i>directional anti-identity</i>	$\neg a \leq b$	WAX	HIGHER
<i>identity</i>	$a \equiv b$	PLATEAU	SAME
<i>anti-identity</i>	$\neg a \equiv b$	NOPLATEAU	DIFF

At a deep level, the only difference between the string-internal and input-output versions of these constraints is the type of correspondence relationship to which they refer. As they are defined here, it appears that the string-internal constraints are different in that they make reference to the linear order of elements within in output string which the input-output constraints do not. This is illusory, the directional asymmetry actually being a product of the properties of correspondence relationships and not an aspect of the formal definition of the constraints. In the following brief and informal set of definitions, this technical niceties and others are ignored in order to give the reader a better idea of what these constraints do in actual analyses:

(6) **String-internal relational constraints**

a. NOWAX[S]

For any corresponding elements α and β in an output where α precedes β and both α and β are in S , β is not higher in S than α . No rises along scale S .

b. WAX[S]

For any corresponding elements α and β in an output where α precedes β and both α and β are in S , β is higher in S than α . Rise along scale S .

c. PLATEAU[S]

For any corresponding elements α and β in an output where both α and β

are in S , α and β are at the same step in S . No changes along scale S .

d. NOPLATEAU[S]

e. PLATEAU[S] For any corresponding elements α and β in an output where both α and β are in S , α and β are not at the same step in S . No plateaus along scale S .

(7) **Input-output relational constraints**

a. NOHIGHER[S]

For any corresponding input-output pair α, β where α is in the input and β is in the output, if both α and β are in S then β is not at a higher step in α in S . No raises along S between input and output.

b. HIGHER[S]

For any corresponding input-output pair α, β where α is in the input and β is in the output, if both α and β are in S then β is not at a higher step in α in S . Must raise along S between input and output.

c. SAME[S]

For any corresponding input-output pair α, β where, if both α and β are in S then β is not at a different step in α in S . No changes along S between input and output.

d. DIFF[S]

For any corresponding input-output pair α, β where, if both α and β are in S then β is not at a different step in α in S . Must change along S between input and output.

The other important constraint set that figures into Structural Optimality analyses consists of constraints on correspondence relationships. These constraints may, de-

pending on the specific constraint, either mandate or penalize correspondence between elements in the output. For a complete treatment of these constraints, see §2.4.

A final note regarding constraints is in order, which is that scale-referring constraints do not replace OT constraints *in toto*. It is still to be assumed that at least MAX, DEP, and markedness constraints are present in Structural Optimality grammars. Such independently motivated constraints will appear in various analyses in this study.

1.1.3 Empirical motivations

The development of this formalism was prompted by the existence of certain phenomena which were difficult to explain otherwise. These phenomena tend to share two characteristics: they are difficult to describe in terms of binary features but easy to define in terms of scale and they make reference, at times, to relations that do not have clear phonetic correlates. Phenomena having these properties include chain shifts (including circular chain shifts), ordering effects in coordinate compounding and reduplication, and certain other phonological effects in reduplication.

1.1.3.1 Chain shifts and circle shifts

Chain shifts are phonological mappings where underlying /a/ corresponds to surface [b] but underlying /b/ corresponds to surface [c]. In rule based phonology, it was easy to model chain shifts since the output mapping could be generated by a sequence of two ordered rules. However, as pointed out by investigators like Foley (1970, 1977), this was not a very insightful way to look at chain shifts since it broke what seemed intuitively to be a single process into a sequence of two processes. That the two rules must be viewed as separate processes is evidenced by the fact that they have independent definitions and must be ordered relative to one another. As Foley demonstrated, the use of phonological scales allowed chain shifts to be treated as single operations. A similar position was taken by Gnanadesikan (1997) in her work on chain shifts, and

Kirchner (1995, 1996) also assumes a scale-like relationship (really, a kind of phonetic grade) in his OT treatment of chain shifts. Thus, it is well established that scales are useful for modelling many types of chain shifts.

None of these earlier models, however, were well suited to model chain shifts which were not monotonic in some phonetic dimension because the scales were assumed to have some phonetic content (though this is less true of Foley's model than the others). A chain shift like that in Eastern A-Hmao is easy to capture through a theory similar to these earlier proposals:

$$(8) \quad H \rightarrow M \rightarrow L$$

However, the cognate chain shift in Shuijingping Hmong is much more problematic (note that $\uparrow H$ means 'super high'):

$$(9) \quad HM \rightarrow \uparrow H \rightarrow H$$

It is not clear what phonetically-grounded scale could be involved in this case. Structural Optimality, however, predicts that exactly this type of chain shift is possible, and even allows these two cognate chain shifts to be modelled in the same fashion.

Just as Structural Optimality predicts the existence of chain shifts of the conventional kind, it also predicts the existence of circular chain shifts (or circle shifts). In the simplest case, a circular chain shift is one where underlying /a/ is mapped to surface [b] and underlying /b/ surfaces as [a]. This is what was known by earlier generative phonologists as an *alpha-switching rule* and is widely called a *toggle*, up to the present. In fact, however, far more complex circle shifts exist, though the known cases are confined to the tone-sandhi systems of Southern Min dialects of Chinese. It has previously been argued that such patterns are marginal and even non-phonological (Tsay and Myers 1996; Moreton 2004b). However, if this study I show that circle shifts are actually more widespread than most earlier investigators have believed, argue that phonological theory should account for them, and show that logical scales can do this very naturally

and elegantly. Furthermore, I show that contrast preservation analyses of these patterns, particularly Barrie (2006), are not adequate to account for the whole range of attested circle shifts.

1.1.3.2 Ordering effects

While constraints that govern the linear order of conjuncts in a coordinate compound (a compound without a single morphological or semantic head) may seem to have little to do with chain shifting alternations, I show that many of the same issues—particularly the issue of scales that do not align neatly with phonetic dimensions—characterize both phenomena. Furthermore, I demonstrate that the same kinds of scales and the same constraints can account for both chain shifts and ordering effects.

Chain shifts appear when HIGHER outranks SAME and SAME outranks ENDMOST. It is HIGHER that drives the unfaithful mappings between input and output. Ordering effects are driven by the string-internal version of HIGHER, which is called WAX. This constraint forces the conjuncts in a compound to be ordered so that there is a rising contour along some scale. For example, in Tangkhul, coordinate compounds are arranged so that higher tonic vowels always come before lower tonic vowels. Thus, conjuncts with the tonic vowel /e/ will always come before those with the tonic vowel /a/:

- (10) a. kə-pí - kə-pem
NOM-sleep NOM-sit
'lodging, etc.'
- b. kə-fək - kə-zà
NOM-drink NOM-eat
'foot/diet, etc.'

The constraint that is active in compelling the ordering is the same constraint (at a fundamental level) as the one that forces a chain shift to occur. This is one of the most surprising and interesting insights of Structural Optimality.

However, as mentioned above, there is another similarity between chain shifts and ordering effects, namely that the scales involved are not necessarily correlated in a monotonic fashion with some phonetic dimension. A good example of this is found in Hmong (Mong Leng). In this language, conjuncts with a high tone and the final syllable are ordered before those with a low tone. Those with a low tone, however, are ordered before those with a mid-tone² This situation is reminiscent of that in the Shuijingping chain shift we examined above. This parallelism provides an example of how Structural Optimality is able to capture the parallelism between phenomena that would otherwise seem completely unrelated.

1.1.3.3 Graded dissimilation and other scalar effects in reduplication

A final empirical motivation for logical scales can be found in certain effects in reduplication constructions. In some cases, as in that of Jingpho, these effects are closely related to the ordering effects that exist in coordinate compounds, but with the special twist that the grammar must chose the right form for the unfaithful conjunct or “reduplication.” In other cases, like that of A-Hmao, the situation is more complicated. In Eastern A-Hmao, it has been observed by Wang and Wang (1996) that there is a three-way patterning of vowels in the nominal reduplication constructions, with one group consisting of the unrounded vowels, the second consisting of the rounded vowels except for /u/, and the final consisting of /u/ alone. I posit a scale consisting of these three categories—unrounded, rounded, and /u/—and show that the generalization governing vocalic alternations in this construction is that the two conjuncts cannot have tonic vowels that are at the same point on this scale. Since the vowel in the first conjunct is always /i/ or /u/, this predicts that the first conjunct has /u/ when the second conjunct has an unrounded vowel, /i/ when the second conjunct has /u/,

²This is a great oversimplification, as will be seen in §5.6, but the point stands and is made even more pointedly in that section.

and can vary between /i/ and /u/ when the vowel in the first conjunct is rounded by is not /u/. This prediction is correct, and the analysis is made possible by the idea that there can be language-specific phonological scales and that there is a class of constraints like NOPLATEAU which penalize corresponding vowels at identical points on a scale. Reduplication constructions, then, also provide evidence for logical scales and the general theory of Structural Optimality.

1.2 Overview of Chapters

This work proceeds from a full introduction to the formal architecture of the theory in Chapter 2 to a presentation of the empirical motivations and applications of the theory in Chapters 3–6. Readers who are not interested in the technical details of the theory may safely skip Chapter 2 and rely on the briefer and less technical presentation in §1.1.2 as their introduction to the formalism that will be used throughout this study. However, beyond its presentation of the constructs and constraints of Structural Optimality, Chapter 2 also includes a demonstration of some of the typological consequences of the theory in §2.7 and readers who skip the earlier sections of Chapter 2 may nevertheless want to read this section.

The next two chapters of the dissertation deal with chain shifts of different types. Chapter 3 shows how logical scales can be used to account for classical chain shifts, taking as examples tonal chain shifts from Western Hmongic languages and consonant-place chain shifts from the speech of children acquiring English. Chapter 4 shows that the same types of constraints and scales account elegantly for circular chain shifts (which have typically been problematic for Optimality Theory and certain other theories of phonology). Furthermore, the chapter shows that such shifts are not so isolated a phenomenon as has been previously believed and that the famous Min tone sandhi circle therefore ought not be dismissed as *sui generis* but should be seen, rather, as but

one example of a robust phenomenon. This fact, I argue, shows that human grammars do not share a property with classical Optimality Theory grammars called *Harmonic Ascent*. Grammars having this property do not allow changes between input and output that do not reduce markedness, and therefore cannot generate circular chain shifts. In the case of both “conventional” chain shifts and circle shifts, I examine the historical factors that lead to the development of these patterns and argue that a true explanation of their properties is to be found in the circumstances that bring them into being rather than some synchronic imperative like contrast preservation.

Having examined patterns of input-output mapping in the form of chain shifts, the study turns, in Chapters 5 and 6, to more output-oriented phenomena. In Chapter 5, I show that the theoretical tools that have already been applied to chain shifts provide the most insightful account for a seemingly unrelated phenomenon, namely ordering effects in coordinate compounds. This Chapter, provides a detailed overview of this little-discussed phenomenon, giving examples from a variety of languages, discusses at some depth the relationship between co-compounds, binomial expressions, and echo reduplication, and shows how facts about these relationships are relevant to the ordering effects under discussion. More relevant to the general theoretical thrust of this work, it shows that an analysis of these ordering effects in terms of Structural Optimality captures a variety of insights about these effects that would be inaccessible otherwise.

The final chapter in the body of this study, Chapter 6, shows that the some otherwise puzzling phonological patterns in reduplication are predicted to exist by Structural Optimality and the theory of logical scales. These include “bounce-back” reduplication in Jingpho and graded-dissimilation in Eastern A-Hmao.

Chapter 2

Logical Scales and the Theory of Structural Optimality

Karma police, arrest this man.
He talks in maths.
He buzzes like a fridge.
He's like a detuned radio.

Thom Yorke, "Karma Police,"
OK Computer

2.1 Structural Optimality

This chapter starts with the notion that grammar is about structure rather than substance (Hjelmslev 1961) and on that foundation outlines a theory of phonological grammar (Structural Optimality) in which patterns and process are not directly motivated by synchronic phonetics but reflect the interaction between abstract representational structures and grammatical constraints¹. It seeks to develop, within the broader framework of Optimality Theory (Prince and Smolensky 1993), an approach to phonology that uses as a primary explanatory device a system of scalar relations and a set of constraints that

¹This is not to say that there is no relationship between phonological form and phonetic substance, but rather that this relationship is largely the result of tendencies in language change rather than being the result of intrinsic properties of a universal grammar and is mediated by "phonetic interpretation". Likewise, Structural Optimality does not rule out the existence of distinctive features which correspond to phonetic contrasts. Indeed, it relies upon the possibility of defining classes of entities, thus assuming the existence of a system of distinctive features. What is novel are the higher-order scalar relationships argued to hold between the entities "composed" of such features.

make reference to them. These scales lack inherent phonetic substance, relating contrasting entities in a paradigm rather than subdividing phonetic space². As such, the theory of Structural Optimality presented here can be seen as a radical response to the phonetic literalism of many (probably most) instantiations of Optimality Theory (see esp. Hayes 1996; Flemming 1995, 2004; Steriade 2002 but also Prince and Smolensky (1993); de Lacy (2002a)).

The first sections of the chapter are devoted to giving formal descriptions of the scales and constraints employed throughout the rest of the dissertation and providing a theoretical rationale for their existence. The second part of the chapter presents a factorial typology of input-output mappings and argues that this typology matches the actual set of such mappings.

2.2 The Definition and Structure of Logical Scales

The heart of the theory presented here is the notion of a phonological scale as a logical—rather than a strictly substantive—entity. These phonological scales are not n -ary features in the same sense as Foley’s (1977) multivariate features or Gnanadesikan’s (1997) ternary scales. They are not simply multivalued distinctive features meant to replace the binary (and privative) features of almost all generative phonology. That sort of feature—the sort that defines a phonological constituent in terms of its substantial properties—might be called an *intrinsic feature*. It is treated, at least in geometric models of phonology, as an entity rather than a property. The scales developed in this work might be classified, instead, as *extrinsic features* in that they classify sets of con-

²As will be seen, by *contrasting entities* we do not mean the set of entities that are in surface or underlying contrast in a particular language. Rather, they represent the set of entities of some type that can be distinguished from one another by the grammar and can potentially be reflected by surface phonetic contrasts.

trasting entities of some type rather than being the substance of which such entities are constructed. That is to say, they are properties rather than entities.

In many respects, this type of scale is to be compared with the sonority hierarchy. All entities of the type “segment” have some position on the sonority hierarchy. However, it has seldom proved useful to introduce sonority as a multivalued distinctive feature (but see Selkirk 1984). Instead, analyses have either sought to derive the sonority hierarchy from more basic binary or privative features (Steriade 1982; Levin 1985; Clements 1990) or to position the sonority hierarchy as an “overlay” that categorizes—rather than defines—the members of the class of segments³. This is precisely the function of the logical scales described here. A difference is to be observed between this class of scales and the sonority hierarchy: while the sonority hierarchy may have one or more consistent phonetic correlates (Parker 2002)⁴, the scales defined here are not required to do so.

Rather than being strictly defined over some phonetic dimension, logical scales are simply orderings over all the contrasting entities of some type. *Ordering*, in the case, is intended in the mathematical sense it is given in set theory, as will be seen later. To see a highly abstract example, take the following elements:

$$(11) \quad \{a, b, c, d\}$$

The possible scales would include all of the following:

$$(12) \quad \text{a. } \{a\} < \{b\} < \{c\} < \{d\}$$

$$\text{b. } \{a, b\} < \{c, d\}$$

$$\text{c. } \{a, b, c\} < \{d\}$$

Speaking more concretely, take the example of a vowel-raising chain shift, a type of alternation that exists quite widely. Given a process that raises underlying low vowels

³For an example of this conception of sonority, see (de Lacy 2002a)

⁴But see also Ohala (1990)

to surface mid vowels but underlying mid vowels to surface high vowels, we could construct the following scale that shows the relative relationships between them:

$$(13) \quad \{a,e,o,i,u\}$$

However, the formally possible scales composed of these vowel segments include all of the following as well:

$$(14) \quad \begin{array}{l} \text{a. } \{i,u\} < \{e,o\} < \{a\} \\ \text{b. } \{a,e,o\} < \{i,u\} \end{array}$$

While it is doubtless true that some scales are more probable (that is, likely to exist in human languages) than others, this consideration is not stipulated in our discussion of their formal structure. All that is required of the scales is that they meet the four conditions that define ordering relations for totally ordered sets, namely reflexivity, antisymmetry (under an equivalence relation), transitivity, and comparability.

Before these conditions are defined, it is important to understand what relations, orders, and ordering relations are. A relation is a set of ordered pairs. The most familiar examples of relations are probably binary operators such as inequalities. For example, \leq is a set of this kind. It is true that $2 \leq 3$ and $2 \leq 2$ but false that $3 \leq 2$ because the set \leq contains the pairs $\langle 2, 3 \rangle$ and $\langle 2, 2 \rangle$ but not the pair $\langle 3, 2 \rangle$. I use the example of \leq because this is the most common name for ordering relations, for reasons that are somewhat beyond the scope of the current study. An order is a pair consisting of a relation and another set. Thus, the ordering of the set of natural numbers (N), could be written as $\langle \leq, N \rangle$. Different types of ordering relations can be described in terms of a number of properties.

Assuming an ordering relation named \leq , the four properties of totally ordered sets (as applied to the scales employed here) can be defined as follows:

$$(15) \quad \text{a. } a \leq a \text{ (reflexivity)}$$

- b. if $a \leq b$ and $b \leq a$ then $a \equiv b$ (antisymmetry)
- c. if $a \leq b$ and $b \leq c$ then $a \leq c$ (transitivity)
- d. $a \leq b$ or $b \leq a$ (comparability)

Reflexivity means that an element stands in the ordering relation (\leq) to itself. To say that the order is antisymmetric, is to say that if a does not stand after b in order and b does not stand after a , then b must be the same as a . This sameness could take two forms. In the natural numbers, it is *identity*: if 5 is less than or equal to x and x is less than or equal to 5, then it holds that x is identical with 5. However, logical scales are not antisymmetric under identity but under *equivalence*. It as if some x could be less than or equal to 5 and 5 could be less than or equal to that x , but that x would not actually be 5, but rather a kind of “5’” that is at the same place in the sequence as 5, but which is not identical to it. The equivalence relationship is represented here with the symbol \equiv . *Transitivity* implies that orders do not “loop back” on themselves. Formally, it means that, given three elements a , b , and c , if a is less than or equal to b and b is less than or equal to c , then a is less than or equal to c . Finally, comparability requires that either a is in the ordering relation to b or b is in the ordering relationship to a , or that any pair of elements within the set can be compared with one another.

A totally ordered set (and thus, one of our scales) is a pair $\langle \leq, T \rangle$ consisting of a set and a relation that orders that set in a manner consistent with the criteria in (15). As will become clear as the theory is developed, all of these properties are indispensable to the evaluation of the constraints that we will define across these scales. It is essential, in the first place, that any element in the set (always consisting of all entities of a particular type) be comparable to itself (reflexivity) and to every other entity in the set (comparability). It is equally important, for reasons that will become clear in Chapters 3 through 6 (as well as in our discussion of directional anti-faithfulness and the constraint ENDMOST), that these orderings (our scales) be transitive, that is, that they contain

no “cycles.” If scales did not have this property, it would be impossible to make the assertion $a \leq b$ in a scale.

The most immediately important of these properties, however, is antisymmetry. This importance derives from the fact that antisymmetry implies the existence of a relation in addition to \leq (the ordering relation) which we will call *equivalence* and which is indicated by the symbol \equiv in (15b). Being equivalent is tantamount to being at the same step on a scale. Note that it is incorrect to define antisymmetry in terms of identity—as is usually done—in this particular case, since it would rule out the possibility of having two non-identical entities that are treated as equivalent within some scale. For example, given the scale:

$$(16) \quad \{a\} < \{e,o\} < \{i,u\}$$

it is essential that e and o behave in equivalent fashion, even though they are not the same entity. It is this relationship that is captured through the equivalence relation. Given the ordering relation with which we started, and this equivalence relation that is a byproduct of the antisymmetry of our ordering, we have all of the relations we need to define constraints over our scales, since the scale referring constraints that we will define are stated directly in terms of these two relations and their negations.

Scales having these properties actually have a different (isomorphic) representation, which is the basis of the notation for scales that will be used in this dissertation. For every total ordering of elements that is antisymmetric under an equivalence relation there is a total ordering of sets that is antisymmetric under an identity relation. In such a representation elements that are equivalent in the first type of ordering are grouped together into sets and it is these sets that are ordered. Thus, given a set

$$(17) \quad \{a,b,c,d\}$$

and the following ordering relation (remembering that ordering relations are sets of ordered pairs, and subsets of the Cartesian product of the set that they order and them-

selves)

$$(18) \quad \{\langle a, a \rangle, \langle a, b \rangle, \langle a, c \rangle, \langle a, d \rangle, \langle b, b \rangle, \langle b, c \rangle, \langle b, d \rangle, \langle c, b \rangle, \langle c, c \rangle, \langle c, d \rangle, \langle d, d \rangle\}$$

implying the equivalence relation

$$(19) \quad \{\langle a, a \rangle, \langle b, b \rangle, \langle b, c \rangle, \langle c, b \rangle, \langle c, c \rangle, \langle d, d \rangle\}$$

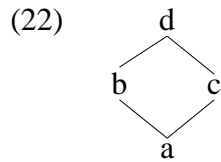
we have an ordering that could also be represented as

$$(20) \quad \left\langle \left\{ \begin{array}{l} \langle \{a\}, \{a\} \rangle, \langle \{a\}, \{b, c\} \rangle, \langle \{a\}, \{d\} \rangle, \\ \langle \{b, c\}, \{b, c\} \rangle, \langle \{b, c\}, \{d\} \rangle, \langle \{d\}, \{d\} \rangle \end{array} \right\}, \{\{a\}, \{b, c\}, \{d\}\}, \right\rangle$$

This representation is an ordering over a partition of the set. Since the ordering is, by definition, transitive, we can eliminate redundant information, and rewrite this ordering as follows (using the novel notation that is employed elsewhere in this dissertation):

$$(21) \quad \{a\} < \{b, c\} < \{d\}$$

It should be noted that that (21) is a notational variant of a Hasse diagram of (20) as shown in (22):



It is important to remember, though, that even though scales will be presented with this more economical notation, the formal definitions of constraints will assume the former representation in which orderings are asymmetrical under equivalence but not necessarily identity.

One convenient aspect of the latter representation of scales, however, is the fact that they provide a convenient indication of how “large” a scale is; that is, how many levels (or sets of equivalent entities) there are in a scale. It should be evident from the definitions given so far, however, that no special importance is attached to this measure.

A scale may be unary, binary, ternary, or n -ary. Evidence will be presented in Chapter 5 for the existence of scales with six or more levels.

These levels exhaustively classify all entities of some type. As the word *type* is used here, it implies a grouping of phonological entities (features, nodes, segments, rhymes, syllables, feet) into class hierarchies. For example, we might say that plosives are a subtype of consonants, and that consonants are a subtype of segments. Any set that is classified by a scale (that is, over which an ordering relation holds) must correspond exactly to some type in this class hierarchy. Of course, just as we can easily imagine the existence of multiple different ordering relations that could order a single set, we must also be able to imagine the existence of multiple scales over the same type. Logical scales are like features in that a single entity may be cross-classified by more than one scale. Just as a vowel may be both [+round] and [-back], even so an entity may occupy the third position on one scale and the first position on another. This point will become important in the discussion of tonal chain shifts and ordering effects.

2.2.1 Natural and Unnatural Scales

It should be abundantly clear from the formal definition given above that there is not a principled way of distinguishing a natural (and typologically likely) scale from an arbitrary scale. Both entities are simply totally ordered sets. However, it is undeniable that most of the phenomena that have been treated as scalar involve some natural phonetic grade. In fact, almost all phonological scales proposed in the literature up to this point divide some phonetic dimension into discrete units. This view of scales, as the direct correlate of gradient phonetics, is much more restrictive than the view expressed above. In fact, it is too restrictive, as I argue in the chapters below. There are a number of phonological phenomena that can only be explained by scales that are *phonetically unnatural*, or even *arbitrary*.

Even granting that such phenomena exist, it still must be explained why natural

scales are more common than unnatural scales (a state of affairs that clearly holds). This problem is part of the much larger naturalness problem in phonology. If phonology does not encode naturalness directly, why is it that most of phonology is phonetically coherent? And if phonology does encode naturalness directly, how do speakers acquire unnatural phonology? I argue for the resolution to the naturalness problem proposed by Ohala (1981, 1993, 1995a,b) and more recently taken up by others (Dolbey and Hansson 1999a,b; Hale and Reiss 2000; Blevins and Garrett 1998, 2004; Blevins 2004). Phonological processes usually have parallels in phonetic processes because phonology is the grammaticalization of phonetics, a process called *phonologization* (Hyman 1977). The avenue of phonologization lies in the domain of learning—misperception and false inferences motivated by a learner’s phonetic knowledge lead to systematic (and phonetically “natural”) differences between her grammar and lexicon and that of earlier cohorts of language learners. Such a framework frees naturalness from Universal Grammar and locates it, instead, in more general (and independently motivated) physical, physiological, and cognitive processes. By the same token, it frees Universal Grammar from the shifting moorings of substance and allows phonology to be a theory of grammar rather than a theory of motor control or psychophysics.

2.2.2 Logical Scales and Richness of the Base

It might be argued that logical scales are incompatible with the Optimality Theory doctrine of Richness of the Base (RotB), seen as a fundamental part of the OT resolution of the duplication problem. The *duplication problem* was a notable embarrassment of early derivational theories of phonology. It was observed by certain practitioners of derivational phonology that many phonological rules duplicated the effects of morpheme structure constraints (that is, MSCs) in maintaining the integrity of certain phonotactic facts (Chomsky and Halle 1968:328; Kenstowicz and Kisseberth 1979, 1977). That is to say, the same phonotactic generalizations were encoded (redundantly)

in both the lexicon and the grammar. Optimality Theory seeks to assign both of these effects to the grammar, and thus remove them entirely from the lexicon. To achieve this end, Optimality Theory stipulates that there be no language specific constraints on underlying form (no MSCs). This aspect of the theory—that there are no arbitrary or language specific constraints on inputs to the grammar—is known as Richness of the Base.

It might seem that the scales defined above make an end-run around RotB by limiting what entities of a particular type the grammar can manipulate: that is, these scales might appear to place language-specific restrictions on possible inputs by only including a subset of the logically possible entities in a given scale. This is an artifact of how the scales have been discussed and exemplified thus far, and not a fact about the theory itself. Given the scale

$$(23) \quad H = \{i,u\} < \{e,o\} < \{a\}$$

one might assume that we have limited the universe of discourse, for vowels, to the five items categorized here. However, the actual scale partitions the whole set of possible vowels. The vowel set given in (23) is minimized only for reasons of notational convenience. This assertion that scales categorize all entities of some type might be problematic if we assume that the set of possible vowels is infinite. However, if phonological entities are made up of a finite number of universal primitives⁵ and if these primitives can only be combined in non-recursive and non-iterative structures, it follows that the set of combinatorial possibilities is finite. That is to say, the number of entities of any type will always be finite, and the inventory of classifiable entities will be the same across languages.

⁵It is assumed here that these primitives are features of the autosegmental type. Other possible formulations, in which these primitives are whole segments, would be conceivable within the same framework, however.

It seems, then, this theory does not present a direct challenge to RotB, but that it does present some very interesting learning challenges. That is, if all possible elements of each type must be organized into scales, and if all scales are constructed, speakers must organize scales with very limited evidence on which those orderings can be based. We will make the assumption that all scales begin their life with a single partition, The ordering relation on each type set begins as the Cartesian product of that set with itself, so for the set T , we would have $\langle T \times T, T \rangle$.

2.3 Structural Harmony Constraints

Having discussed the formal structure of phonological scales, we will now introduce the mechanism by which the rest of the grammar interacts with such scales. Optimality Theory, in its conventional formulation, consists of constraints over outputs—structural harmony constraints or (to use the more common term) “markedness” constraints and constraints over input-output correspondences (“faithfulness” constraints). The approach argued for here employs markedness-like constraints that can make reference to scales. Similar claims have been made by Gnanadesikan (1997) and de Lacy (2002a).

I will propose three families of scale-referring structural harmony constraints: ENDMOST, TOP, and EXTREME. By far, the most important of these is ENDMOST. However, the other constraints appear in certain of the following analyses so they are given definitions here as well. In order to make the definitions of the constraints more concise, I introduce a number of definitions here:

- (24)
- a. T is the set of all elements of some type.
 - b. S is the totally ordered set $\langle \leq, T \rangle$.
 - c. \equiv is an equivalence relation such that $a \leq b$ entails $b \leq a$.

- d. $<$ is a total order, antisymmetric under identity, over equivalence classes defined by \equiv over T such that $A, B \in E$, $\alpha \in A$, $\beta \in B$, and $A < B$ entails that $\alpha \leq \beta$ and not $\beta \leq \alpha$ in S .
- e. \mathbb{C} is an output candidate.

2.3.1 ENDMOST

I propose a family of scale-referring structural harmony constraints called ENDMOST⁶, which is sometimes abbreviated, in this work, as END. This constraint penalizes segments (or other entities) that are classified by a scale to the extent that they are not at the far (“unmarked”) end of the scale. Informally, we will treat it as if it assesses one violation for each increment that lies between the entity in question and the end of the scale. Formally, it may be defined as in (25):

(25) ENDMOST[S]

There is no equivalence class A in S/\equiv for which there is some $\alpha \in \mathbb{C}, A$ and some equivalence class B in S/\equiv where $B < A$ in E .

(26) ENDMOST[S] (informal)

There is no step on the scale between an entity in the scale S and the bottom of the scale.

This constraint is formulated specifically so as to avoid reference to gradient violability or cardinality. What the constraint penalizes are steps in a scale—sets in a partition—that sit between entities belonging to the scale and the lowest step (partition) in the scale, rather than concrete entities in the output string. In this formulation, this constraint is no more gradiently violable than any other constraint that could be violated

⁶This constraint is not meant to replace all other “markedness” constraints. Indeed, the theoretical apparatus introduced here assumes the whole standard cohort of constraints on prosodic structure, to name but one set of examples. ENDMOST, though, is posited as the only scale referring structural harmony constraint.

at multiple points within a single candidate, and does not require the grammar to count or perform arithmetic any more than these other constraints. The counting problem remains an issue within OT, despite the fact that it was addressed as early as Prince and Smolensky (1993), since the formalism seems to require the grammar to compare the length of lists of violations, and while this may not be an arithmetic operation in and of itself, it appears that the computational mechanisms that are needed to compute this relation either include arithmetic operations or are sufficient to construct computational mechanisms capable of counting and arithmetic. This is not a problem that can be solved here, however, and it is sufficient to note that ENDMOST is not gradiently violable in any special sense, even though it will be treated that way in subsequent discussion (wholly to ease the exposition of ranking arguments).

ENDMOST could equally well be decomposed into a series of constraints referencing specific points along a scale. This could be achieved most elegantly by decomposing ENDMOST constraints into sets of “stringent” scale referring constraints of the type detailed in de Lacy (2002a). For example, if there was a scale:

$$(27) \quad S = \{i,u\} < \{e,o\} < \{a\}$$

instead of the constraint ENDMOST[S], we could propose the following constraints, which would have the same effect, except when they were “inter-ranked” with other constraints:

- $$(28) \quad \begin{array}{l} \text{a. } *a \\ \text{b. } *a/e/o \\ \text{c. } *a/e/o/i/u \end{array}$$

This formulation would have three disadvantages: **(1)** it would make the theory slightly less restrictive by allowing the scale-referring markedness constraints to be inter-ranked with other constraints, **(2)** it would require constraints to make specific reference to points along the scale, rather than to relative positions along the scale and **(3)** it would

make the notational description of analyses considerably more complex. However, all of the analyses given here would work equally well if ENDMOST was atomized into a series of smaller constraints.

2.3.2 TOP

In addition to ENDMOST, we must posit two other scale-referring structural harmony constraints. The first of these is TOP, which penalizes elements that are the top of the relevant scale. Contrary to first appearance, TOP is not simply the inverse of ENDMOST; while fewer violations of ENDMOST are incurred by candidates that are near the bottom end of the scale than by candidates that are near the top end of the scale, TOP is only sensitive to whether or not a candidate is at the highest step on the scale.

(29) TOP[S]

There is no $\alpha \in \mathbb{C}$ such that A, B are equivalence classes in S/\equiv , $\alpha \in A$, and $A < B$ in E .

(30) TOP[S] (informal)

An entity in the scale S must be at the top of the scale.

2.3.3 EXTREME

The last of the Structural Harmony constraints posited in this study is EXTREME, which penalizes candidates that are not at the extreme ends of a scale (or, in other words, those that lie in the middle). It is defined as follows:

(31) EXTREME[S]

There is no α in \mathbb{C} for which there are three (distinct) equivalence classes $A, B, C \in S/\equiv$ such that $\alpha \in A$, $B < A$, and $A < C$ in E .

(32) EXTREME[*S*] (informal)

There is no entity at a step on the scale *S* with both a step above it and a step below it.

2.4 String Internal Correspondence

The theory described here rests upon the theories of string-internal (syntagmatic) correspondence that are developed by Walker (2000a), Hansson (2001), and Rose and Walker (2004) (among others). While this theory does not directly relate to the theory of phonological scales, it is essential to the definition of scale-referring correspondence constraints. Specifically, in order to reduce the paradigmatic and syntagmatic (anti-) identity constraints to a single set, it is necessary to allow correspondence relationships of this type. This current in the ocean of Optimality Theory comes with two major claims:

1. Just as there are correspondence relationships between inputs and outputs, there are also correspondence relationships between entities in the output.
2. Unlike input-output correspondence, this syntagmatic correspondence is governed by violable constraints that are part of the ordinary OT constraint hierarchy. These constraints consist of one hierarchy that enforces correspondence between elements of various degrees of similarity and another hierarchy that penalizes correspondence relations between entities that are not in close proximity to one another.

As developed by Walker (2000a), Hansson (2001), and Rose and Walker (2004), the theory of string internal correspondence was focused primarily upon long-distance relationships between consonants. To this end, Rose and Walker (2004) propose a similarity-based hierarchy correspondence constraints as in (33):

(33) Similarity-based correspondence hierarchy (consonants)

$\text{CORR-T} \leftrightarrow \text{T} \gg \text{CORR-T} \leftrightarrow \text{D} \gg \text{CORR-K} \leftrightarrow \text{T} \gg \text{CORR-K} \leftrightarrow \text{D}$
'identical stops' 'same place' 'same voicing' 'any oral stops'

(Adapted from Rose and Walker 2004:491)

These constraints compel similar consonants to enter into correspondence relationships with one another. They interact with faithfulness constraints that hold over consonants in the output (e.g. IDENT-CC), input-output faithfulness constraints, and PROXIMITY constraints, in determining whether output consonants are in correspondence with one another in the winning candidate. This later class of constraints, PROXIMITY, deserves special comment. While Hansson (2001) factors a proximity hierarchy into his hierarchy of correspondence constraints, yielding a large number of constraints that refer both to similarity and proximity, Rose and Walker (2004) posit a single constraint

(34) PROXIMITY

Correspondent segments are located in adjacent syllables.

The ranking of this constraint relative to the similarity hierarchy determines which types of correspondence relationships will obey proximity and what types of correspondence relationships can exist beyond adjacent syllables. This constraint will be assumed in the analyses given here. However, at least for purposes of demonstration, it will be necessary to posit an additional constraint that militates against any correspondence relationships at all, even those between entities in adjacent syllables. We may call this constraint NOCORR:

(35) NOCORR

Let \mathbb{C} be an output candidate; if $\alpha, \beta \in \mathbb{C}$ then $\neg(\alpha \leftarrow \beta)$; that is, α and β are not correspondents of one another.

Some significant additions have to be made to the theoretical machinery constructed by Rose and Walker (2004). In particular, the case studies described here involve corre-

spondence relationships between rhymes, vowels and tones rather than between consonants. Furthermore, the similarity comparisons that need to be made cannot be derived simply from the featural make-up of the entities under comparison⁷ The similarity relationships that establish correspondence relationships in several of the case studies presented here are based not simply upon intrinsic featural content but rather upon positional factors. Specifically, it can be observed that vowels and tones in prosodic head positions preferentially enter long-distance correspondence relationships with one another.

One way of formalizing this observation is to allow correspondence relationships between two syllables rather than forcing correspondence relationships to hold directly between individual segments. If two syllables were in correspondence, each of their respective subconstituents would be in correspondence with one another through a kind of CORRESPONDENCE BY INHERITANCE. Through this type of mechanism, it would not be necessary to devise correspondence constraints that referred, for example, to tones born by stressed syllables. We could give the relevant constraints the following definition:

$$(36) \text{ CORR-}\sigma \leftarrow \sigma/P$$

Let σ_1 and σ_2 be syllables in a syllabified output string \mathbb{C} ; if σ_1 , and σ_2 are in P , the set of all syllables in some structurally prominent position and $\sigma_1 \prec \sigma_2$ in \mathbb{C} then $\sigma_2 \leftarrow \sigma_1$; that is, σ_1 (and its subconstituents) are in a correspondence relationship with, or depend upon, σ_2 (and its subconstituents).

⁷Of course, the similarity hierarchy that Rose and Walker (2004) propose cannot be derived directly from a measure of shared features. Rather, it attaches a relative priority to two kinds of featural relationships: place of articulation and voicing (in that order). However, the hierarchy proposed by Rose and Walker is different from the similarity relationships described here in that Rose and Walker's hierarchy makes no reference to the context in which the consonants occur, but only to facts about their intrinsic content.

In informal terms, this may be restated as follows:

- (37) If there are two syllables in an output candidate, one preceding the other, and if they are both in the same, specific, prominent structural position (head of the foot, beginning of a PrWd, and the like) then the first should correspond to (depend upon) the second.

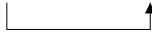
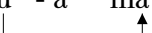
Given this constraint family, we could posit the following similarity-based correspondence hierarchy:

- (38) Similarity-based correspondence hierarchy (syllables)

$$\begin{array}{ccc} \text{CORR-}\acute{\sigma} \leftarrow \acute{\sigma} & \gg & \text{CORR-}\sigma \leftarrow \sigma \\ \text{'stressed syllables'} & & \text{'any syllables'} \end{array}$$

The intuition captured in this hierarchy are quite simple, if somewhat paradoxical: *ceteris paribus*, two stressed syllables are more similar to one another than two two unstressed syllables. Further, unstressed syllables are no more similar to one another than one stressed and one unstressed syllable. The purpose of this stipulation is to account for the fact that correspondence relationships between vowels or tones in prominent positions appear to be able—indeed, prone—to skip over non-prominent entities that lie between them. An example is vowel disharmony in A-Hmao echo reduplication. In this language construction there is a difference requirement that holds only between the tonic vowels of the two conjuncts (skipping, and not applying to, the unstressed vowels that lie in between)⁸:

⁸It might be noted that under a base-reduplicant correspondence theory (McCarthy and Prince 1995) analysis of this construction, the disharmonic vowels would already be in correspondence. Given a rather idiosyncratic notion of positional faithfulness, it may be possible to generate this pattern in that model, without a special correspondence relationship between the final vowels of the two conjuncts. However, as will be seen in Chapter 5 below, the analogous ordering effects in coordinate compounds require an identical type of relationship which cannot be reduced to base-reduplicant correspondence.

- (39) a. $a^{H_{nd}ú^{HM}}$ ‘edge’ $a^{H_{nd}í^{H-}}$ $a^{ML_{nd}ú^{ML}}$ ‘thereabouts’

 b. $a^{H_{m}á^{HM}}$ ‘eye’ $a^{H_{m}ú^{H-}}$ $a^{ML_{m}á^{ML}}$ ‘eyes, etc.’


The ranking that generates this kind of correspondence relationship would be one in which $CORR-\acute{\sigma} \leftarrow \acute{\sigma}$ dominates NOCORR which dominates $CORR-\sigma \leftarrow \sigma$:

- (40) $CORR-\acute{\sigma} \leftarrow \acute{\sigma} \gg NOCORR \gg CORR-\sigma \leftarrow \sigma$

2.4.1 Properties of correspondence relationships

It is important to note that correspondence is not simply a coindexation relationship. Coindexation relationships are transitive: if some entity α is coindexed with another entity β , and β is coindexed with γ , it follows that α is coindexed with γ . As should be evident to anyone with a passing familiarity with base-reduplicant correspondence theory (BRCT) or with various lexical phonology optimality theory models (LPM-OT), correspondence relationships are not transitive in this respect. We can state this more formally as in (41):

- (41) NON-TRANSITIVITY OF CORRESPONDENCE

Correspondence relationships are not transitive: $(\alpha \leftarrow \beta \wedge \beta \leftarrow \gamma) \not\Rightarrow \alpha \leftarrow \gamma$.

That α is in a correspondence relationship with β and β is in a correspondence relationship with γ does not entail that α is in a correspondence relationship with γ .

We must view correspondence, then, as a dependency relationship rather than a relationship like coindexation.

It is also essential that correspondence relationships be asymmetrical. As Hansson (2001) notes, this move might seem to be ad hoc, but it is actually completely necessary to OT as it is currently practiced (that is, in its Correspondence Theory incarnation).

This same point was raised earlier by Archangeli (1999) (and, in a different way, by Walker (2000b)). As both Archangeli (1999) and Hansson (2001) demonstrate, the Optimality Theory constraints MAX and DEP are indistinguishable without asymmetrical correspondence. This fact is concealed somewhat in input-output relationships, where asymmetry is implicit (inputs influence outputs; outputs do not influence inputs, except perhaps in lexicon optimization). In correspondence relationships within strings, however, this becomes more obvious. If one can posit constraints of the MAX-BR type, which penalizes reduplicants which lack some structure found in the input, but does not penalize reduplicants which have some structure not found in the input, correspondence relationships must be understood to be asymmetrical.

Hansson (2001) makes the claim that correspondence relationships, at least within consonant harmony, are right-to-left. This is consistent with a number of details of consonant harmony, and Hansson makes an strong case for this restriction in correspondence relationships. I will not make this assumption, and will instead treat correspondence relationships as having the opposite directionality (left to right). However, it should be noted that this decision has been made purely for aesthetic reasons and that an identical analysis of the cases examined here could be formulated with the opposite directionality.

2.5 Syntagmatic Relational Harmony Constraints


Given a notion of string-internal correspondence, it is possible to define sets of constraints over elements in such a relationship. These constraints allow the theory to account for phenomena such as assimilation, dissimilation, and ordering effects. Many of these constraint types are already discussed in the literature, as will be seen below. What makes these constraints novel—and what justifies their novel monikers—is that they are evaluated relative to phonological scales.

It will become evident, as the exposition proceeds, that these constraints are completely parallel to the input-output constraints described in Section 2.6. In essence, I am not proposing six novel constraints. Rather, I am proposing three novel constraints, the only significant difference between the two incarnations of each constraint being that the output-oriented constraints make reference to linear precedence where the input-output correspondence constraints make reference to level precedence (input preceding output). A major strength of this theory is its ability to predict a correct typology of both input-output mappings and scalar relationship across strings using constraint sets that are completely isomorphic.

All of the correspondence constraints described here are categorical. That is to say, a single pair of corresponding entities can only violate a constraint once, regardless of how ill-formed they are relative to that constraint. A constraint of this type does not count violations in terms of number the of scale steps between the actual pair and the pair that would satisfy the constraint (as ENDMOST superficially appears to do). For example, suppose there is a scale

$$(42) \quad H = \{a\} < \{e\} < \{i\}$$

and a constraint PLATEAU-H, which says that members of an input-output pair should have the value along H . Under this state of affairs, the input-output pair $\langle a, i \rangle$ would have the same violation profile as $\langle a, e \rangle$, that is to say, both pairs would incur one violation of the constraint. In explaining the constraints below, I will assume the scale over the vowels given in (42) and will use examples in the following (contrived) format:

$$(43) \quad \text{əp}^{\text{h}}\text{á kəp}^{\text{h}}\text{í}$$


2.5.1 PLATEAU

PLATEAU is essentially a scale-referring version of the constraint AGREE (Lombardi 1999; Baković 2000). As such, it is functionally equivalent to the ASSIM constraint of Gnanadesikan (1997). Hansson (2001) insightfully notes that constraints of this type are essentially syntagmatic analogues to the IDENT constraints of classical Optimality Theory.

Informally defined, PLATEAU states that segments or other constituents, that are in a correspondence relationship within a string must be at the same point along some scale, assuming that they are both classified by that scale. This constraint may be defined more explicitly as follows:

(44) PLATEAU[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation on an output string \mathbb{C} ; for any $\alpha, \beta \in \mathbb{C}$ if $\alpha, \beta \in T$ and $\alpha \leftarrow \beta$ then $\alpha \equiv \beta$, that is, $\alpha \leq \beta$ and $\beta \leq \alpha$.

Less formally, it may be defined:

(45) PLATEAU[S] (informal)

Corresponding elements are at the same point on the scale S .

Given the vowel scale in 42, the following candidates (among others) would satisfy PLATEAU:

(46) a. əp^há kəp^há b. əp^hí kəp^hí c. əp^hú kəp^hí d. əp^hé kəp^hó

Candidates that would violate PLATEAU include those below:

(47) a. əp^há kəp^hí b. əp^hé kəp^hí c. əp^hó kəp^há d. əp^hí kəp^hé

PLATEAU may be motivated by the same set of phenomena that have been captured previously through syntagmatic agreement constraints, namely various types of assimilation (both local and non-local). The heavy work of establishing correspondence relationships between entities in the output is performed by correspondence constraints as described in Section 2.4.

2.5.2 NOPLATEAU

NOPLATEAU represents the side of the correspondence coin opposite PLATEAU. It penalizes candidates in which corresponding entities in the output have the same scale value. Among other things, this constraint motivates dissimilations, both the local and the non-local variety. Given the scale we formulated above in 42, the following candidates would satisfy NOPLATEAU:

- (48) a. əp^há kəp^hí b. əp^hé kəp^hí c. əp^hó kəp^há d. əp^hí kəp^hé
-

However, these candidates would satisfy NOPLATEAU:

- (49) a. əp^há kəp^há b. əp^hí kəp^hí c. əp^hú kəp^hí d. əp^hé kəp^hó
-

It can be formally defined as in (50):

$$(50) \text{ NOPLATEAU}[S]$$

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation on an output string \mathbb{C} ; for any $\alpha, \beta \in \mathbb{C}$ if $\alpha, \beta \in T$ and $\alpha \leftarrow \beta$ then $\neg(\alpha \equiv \beta)$, that is, $\neg\alpha \leq \beta$ or $\neg\beta \leq \alpha$.

Informally, the following definition will suffice:

(51) NOPLATEAU[S] (informal)

Corresponding elements are not at the same point on the scale *S*.

The exact parallelism between this constraint and PLATEAU makes an interesting prediction, namely that all cases of assimilation have a dissimilatory counterpart. The same claim has been made earlier by Hyman (2000). This hypothesis has yet to be tested fully, but currently appears both promising and problematic. For example, it is relatively easy to think of cases of voicing or place assimilation between contiguous consonants but relatively difficult to think of examples of the opposite type of processes (place dissimilation or voicing dissimilation between contiguous consonants). Voicing dissimilation does occur, but it is typically a long-distance (rather than a local) effect. Likewise, it is relatively easy to think of cases of spreading vowel harmony, but relatively difficult to think of examples of spreading disharmony. Cases of vowel disharmony are typically limited to echo reduplication constructions⁹.

Perhaps the best resolution to this problem is not to discard NOPLATEAU which is, as will be shown below, necessary to account for certain phonological phenomena, but rather to attribute this disparity to factors lying outside of the grammar. It has been claimed by Ohala (1981) and others that there are differences in the factors that give rise to assimilation and dissimilation historically. It seems very likely that speakers are perfectly capable of learning dissimilation patterns (Pycha et al. 2003) but that history is relatively unlikely to present speakers with the opportunity. Vowel harmony, for example, may arise from the misinterpretation by speakers of vowel-vowel coarticulation (Ohala 1994b,a; Przedziecki 2005)¹⁰. Vowel disharmony, on the

⁹There are counter-examples to this generalization. Two of the best documented cases of vowel dissimilation are the low vowel dissimilation found in certain Oceanic languages (Blust 1996a,b; Lynch 2004) and the backness dissimilation found in Yucatec Maya (Krämer 2000)

¹⁰This hypothesis is not without its problems and clearly cannot account for all aspects of vowel harmony (e.g. unbounded spreading, certain biases in directionality). However, it does provide a con-

other hand, would have to arise from some other mechanism—perhaps stylistic devices that require metalinguistic knowledge.

2.5.3 WAX and NOWAX

Both PLATEAU and NOPLATEAU have analogues in constraint families that have been proposed by earlier investigators. I will now introduce two new constraints, WAX and NOWAX. WAX is like NOPLATEAU in that it enforces a difference requirement between corresponding elements; NOWAX is like PLATEAU in that it penalizes certain types of differences between corresponding elements. It will become clear how this is true shortly. Though they share these similarities with their symmetrical cousins, they are actually simpler than PLATEAU and NOPLATEAU in that they consist of nothing more than the ordering relation we have called \leq and its negation. That is to say, they are the most primitive constraints that could be stated in terms of our scales—they can be constructed wholly without logical conjunction.

WAX and NOWAX appear to be different from the PLATEAU constraints in that they make use of linear precedence relationships. However, when the matter is examined more closely, they rely on the asymmetry that is inherent in the correspondence relationship described in §2.4 WAX says that there must be a difference between two corresponding elements along some scale. NOWAX penalizes such a difference. This is not accomplished by adding some additional technology to the theory, but rather by stripping away, for the sake of these constraints, some of the technology we have needed elsewhere. Seen in this light, it is actually the more normative-looking convincing account for the genesis of vowel-vowel assimilation. An example from the authors own research concerns Tankghulic languages. In Standard Tankghul, there is no vowel harmony but there is very significant anticipatory coarticulation between the vowels in unstressed prefixes and the vowels in the following stressed roots. In East Tusom, in this same context, a nascent system of vowel harmony has developed.

straints PLATEAU and NOPLATEAU which require additional stipulation, and not the directional constraints NOWAX and WAX.

NOWAX is defined, quite simply, as follows:

(52) NOWAX[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation on an output string \mathbb{C} ; for any $\alpha, \beta \in \mathbb{C}$ if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\alpha \leq \beta$.

By way of illustration, the following candidates would satisfy NOWAX:

(53) a. əp^há kəp^há b. əp^hí kəp^hú c. əp^hú kəp^hó d. əp^hó kəp^há

These candidates, however, would violate NOWAX

(54) a. əp^há kəp^hé b. əp^hé kəp^hí c. əp^há kəp^hí d. əp^hé kəp^hó

Its counterpart, WAX, is simply the negation of this constraint:

(55) WAX[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation on an output string \mathbb{C} ; for any $\alpha, \beta \in \mathbb{C}$ if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\neg \alpha \leq \beta$.

Informally, WAX is even simpler

(56) WAX[S] (informal)

If there are two elements α and β in the scale S , and α corresponds to β then β is higher on S than α .

The violation profile of WAX is opposite that of NOWAX (since it is the negation of NOWAX), so the candidates in (54) satisfy this constraint while those in (53) violate it. This constraint family is motivated by the existence of “cline effects” in the ordering of coordinate compounds, where two stems in a word are ordered so that some characteristic of the first is “lower” along some scale than the corresponding characteristic in the second. Take, for example, Jingpho, where coordinate compounds are ordered so that the tonic vowel of the second stem is more sonorous than the tonic vowel of the second stem if this is possible without changing the vowel quality in either stem (Dai 1990a). In this case, we would have the “sonority” scale S :

$$(57) \quad S = \{i,u\} < \{e,o\} < \{a\}$$

And the constraint WAX-S:

$$(58) \quad \text{WAX-S}$$

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation on an output string \mathbb{C} ; for any $\alpha, \beta \in \mathbb{C}$ if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\neg \alpha \leq \beta$.

The effect would be to penalize candidates in which the stems were linearized in order of decreasing sonority. To see how this functions in concrete terms, see §5.5.

There are two logical ways of accounting for ordering phenomena of this kind. One is through a constraint of this type (a constraint that favors outputs where there is a directional cline along some dimension). The other approach is to posit an asymmetry between the positions in which the two conjuncts could be ordered (a head versus a non-head) and posit markedness constraints that favor (for example) high-sonority vowels more than low sonority vowels in head positions. As will be seen below, this second approach is inferior to the first both in that it requires the linguist to posit the existence of prosodic structure for which there is no other evidence.

WAX is the syntagmatic analogue of the input-output constraint HIGHER, which will be discussed below in Section 2.6.3.

2.6 Input-Output Relational Harmony Constraints

The same set of scale-referring constraints that can be motivated over syntagmatic correspondence relations must also be posited for input-output relations in order to model the set of attested input-output mappings. They are given different names only as a matter of convenience. The relationship between the two sets of names is given in (59):

	<i>String Internal</i>	<i>Input-Output</i>	<i>Description</i>
	PLATEAU	SAME	general identity
(59)	*PLATEAU	DIFF	general anti-identity
	*WAX	NOHIGHER	directional identity
	WAX	HIGHER	directional anti-identity

2.6.1 SAME

SAME is the input-output version of PLATEAU, and both of these constraints are scale-referring versions of the standard OT constraint IDENT. Its formal definition is as in (60), and—as can be seen—is formally identical to the definition of PLATEAU given above.

(60) SAME[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation from an output candidate \mathbb{C} to an input \mathbb{I} ; for any $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{I}$, if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\alpha \equiv \beta$, that is $\alpha \leq \beta$ and $\beta \leq \alpha$.

This family of constraints is the most obvious extension of phonological scales to correspondence theory. It does not significantly increase the formal power of the theory, since it is simply an extension of an already existing mechanism to a subtly different

type of representation. By way of exemplification, the candidates in (61) satisfy SAME while those in (62) do not:

(61) a. $\langle ka_i p, ka_i p \rangle$ b. $\langle ke_i p, ke_i p \rangle$ c. $\langle ke_i p, ko_i p \rangle$ d. $\langle ki_i p, ku_i p \rangle$

(62) a. $\langle ka_i p, ke_i p \rangle$ b. $\langle ke_i p, ki_i p \rangle$ c. $\langle ko_i p, ku_i p \rangle$ d. $\langle ka_i p, ke_i p \rangle$

2.6.2 DIFF

The constraint family DIFF instantiates a far more controversial idea that constraints like PLATEAU and SAME. It enforces a general anti-identity requirement over corresponding pairs in the same manner as NOPLATEAU. It is defined as in (63):

(63) DIFF[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation from an output candidate \mathbb{C} to an input \mathbb{I} ; for any $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{I}$, if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\neg(\alpha \equiv \beta)$.

The violation profile of DIFF is exactly the opposite of that for SAME, and its violation profile is, therefore, opposite that of SAME such that the candidates in 62 satisfy DIFF and those in (61) violate DIFF.

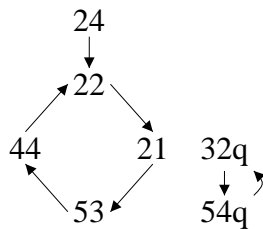
Anti-identity (Alderete 2001), as exemplified by DIFF, is controversial partly because it undermines one of the most interesting formal results in Optimality Theory, namely the principle of Harmonic Ascent. Moreton (2004b) provided a formal proof that an optimality theoretic grammar consisting wholly of markedness and faithfulness constraints (a CLASSICAL OT grammar) can generate any mapping except an infinite chain shift and a circular chain shift. Circular chain shifts are impossible because, in Classical OT, any difference between input and output must improve a candidate relative to markedness. Suppose that some OT grammar maps input A to output B. It follows that B must be less marked, in some dimension, than A. It is impossible for any

such grammar to map B to A since A is different from B and B is less marked than A. Thus, a circle shift in which A is mapped to B but B is mapped to A is predicted to be impossible. This principle is called Harmonic Ascent.

It is hard to argue with the formal beauty of Moreton's Theorem, or the attractiveness of Harmonic Ascent. There is considerable evidence, however, that Harmonic Ascent is merely a property of Classical Optimality Theoretic grammars and not of human languages.

Moreton (2004b) addresses the best-known counter-example to the general claim that circular chain shifts do not exist outside of morphology, namely the Southern Min tone sandhi circles. Take, for example, the famous instance from Xiamen/Mainstream Taiwanese, where arrows point from inputs to the outputs that surface in sandhi context (Wang 1967; Cheng 1983; Ballard 1988; Chen 2000; Moreton 2004b):

(64) **The Xiamen/Mainstream Taiwanese tone sandhi circle**



Moreton concludes that all phonological explanations that have been offered for this phenomenon are ad hoc and offer no real insight into its true nature. He argues, further, that the Taiwanese tone circle is not necessarily a counter-example Harmonic Ascent because it involves an interaction between syntax and phonology. This explanation is not wholly satisfying, in particular because there are other cases of circular tone shifts that do not appear to be syntactically conditioned (Wang and Wang 1986; Dai 1990c; Johnson 1999; Mortensen 2003). Furthermore, excluding all phonology that is sensitive to morphology and syntax from our search for the range of possible phonological grammars would radically alter the endeavor, limiting the admissible evidence, for the

most part, to relatively low-level automatic processes.

This discussion is relevant because admitting any anti-identity input-output constraint into the phonological grammar undermines Harmonic Ascent completely. In the string-internal domain, in contrast, this is not a problem; a constraint like NOPLATEAU could never produce a circular chain shift, any more than a markedness constraint could. It is the cross-level nature of DIFF that undermines Harmonic Ascent. Given the same scenario as above, even if A is more marked than B, B can still be mapped to A if DIFF (or an equivalent anti-identity constraint) is ranked above markedness.

Anti-identity is not the only mechanism proposed within Optimality Theory that undermines Harmonic Ascent. Perhaps the best-known such proposal is McCarthy's (2002) Comparative Markedness. In this theory, markedness constraints are decomposed into matching pairs of (two level) constraints: one member of each pair counts new violations (violations that are not present in the fully faithful candidate or FFC) while the other member counts old violations (violations that are present in the FFC). This theory is able to account for various types of phonological opacity including derived environment effects, counterfeeding and counterbleeding. McCarthy points out that this framework can generate both infinite chain shifts and circular shifts. For example, when old markedness is ranked above faithfulness which is, in turn, ranked above new markedness, a circular chain can result, as demonstrated in (65) and (66) (adapted from (McCarthy 2002)):

		/e/	O*HIGH	O*MID	IDENT(high)	N*HIGH	N*MID
(65)	➡ (a) i					*	
	(b) e			*!			
		/i/	O*HIGH	O*MID	IDENT(high)	N*HIGH	N*MID
(66)	(a) i		*!				
	➡ (b) e				*		*

McCarthy considers this to be a major problem with comparative markedness. As will be shown below, this may actually be a strength of comparative markedness theory, since circular chain shifts are attested—a strength that it shares with the theory described here.

Comparative markedness grammars have the potential to do something that grammars containing DIFF cannot do: to generate infinite chain shifts. As is noted by McCarthy (2002), it is possible to motivate unconditioned epenthesis with comparative markedness constraints. Given a grammar with an undominated ONSET and NOCODA and MAX ranked above DEP, if $O_{FINAL-C}$ dominates DEP which dominates $N_{FINAL-C}$, then unconditioned epenthesis can result, as demonstrated in (67) (adapted from (McCarthy 2002)):

		/pata/	$O_{FINAL-C}$	DEP	$N_{FINAL-C}$
(67) a.	☞ (a)	pata?i		**	*
	(b)	pata	*!		
		/pata?i/	$O_{FINAL-C}$	DEP	$N_{FINAL-C}$
b.	☞ (a)	pata?i?i		**	*
	(b)	pata?i	*!		

In (67), epenthesis occurs because it is always better to have a new vowel at the end of a word than an old vowel (which would violate the higher ranked $O_{FINAL-C}$). Since some segment has to be epenthesized, and since syllables have to have a CV shape, a full syllable will always be epenthesized, regardless of how long the input string is.

A major virtue of a theory containing DIFF is that it can generate a circular chain shift (of a certain type) without being able to generate an infinite chain shift (without additional theoretical machinery). DIFF is incapable of motivating epenthesis as epenthesis, and it therefore cannot generate unconditioned epenthesis. It should be noted that one could mimic epenthesis in a grammar containing DIFF by defining a

scale over a set of strings such that each string was one segment longer than the string at the preceding step. Indeed, under the same set of assumptions, it would be possible to define a scale over a set of unrelated strings, allowing for a completely random chain shift. However, for independent reasons, scales are defined as partitions of all entities of a given type, with the further (logical) qualification that these entities are neither recursive nor sequential (a necessary stipulation if the set of entities classified by a scale is to be finite). Strings are sequential structures, and are not, therefore, classified by scales. As a result, the theory cannot generate infinite chain shifts.

DIFF by itself (or rather, in interaction with standard markedness constraints and the scale referring constraints we have discussed thus far) can only generate a very limited set of chain shifts, all of which contain circles. Such circles are always minimal; they are “exchanges” between the “least marked” and “second least marked” points on the scale. Generating longer chain shifts, including longer circular shifts requires additional technology.

2.6.3 HIGHER and NOHIGHER

This technology is already implicit in the constraints that have been defined so far. Among the syntagmatic correspondence constraints, I introduced the general “dissimilation” constraint family NOPLATEAU and the directed dissimilation constraint WAX. It is reasonable that the kind of correspondence relationship captured in WAX should exist between input-output correspondence pairs as well. This analogy is somewhat trickier than the analogy between PLATEAU and SAME or NOPLATEAU and DIFF since HIGHER makes reference to a linear precedence relation which does not appear to exist across input-output pairs. The solution adopted here is to translate the linear precedence relationship into a “level-precedence” relationship such that the input “precedes” the output. This notion that the input-output relationship is parallel to a left-to-right precedence relationship syntagmatically is not entirely novel and could be seen as a natural

outgrowth. Applied to WAX, we are able to infer the existence of the constraint family HIGHER:

(68) HIGHER[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation from an output candidate \mathbb{C} to an input \mathbb{I} ; for any $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{I}$, if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\neg(\alpha \leq \beta)$.

The candidates in 69 satisfy HIGHER, while those in 70 violate it:

(69) a. $\langle ka_{ip}, ke_{ip} \rangle$ b. $\langle ka_{ip}, ko_{ip} \rangle$ c. $\langle ke_{ip}, ki_{ip} \rangle$ d. $\langle ko_{ip}, ku_{ip} \rangle$

(70) a. $\langle ka_{ip}, ka_{ip} \rangle$ b. $\langle ke_{ip}, ko_{ip} \rangle$ c. $\langle ko_{ip}, ka_{ip} \rangle$ d. $\langle ki_{ip}, ke_{ip} \rangle$

HIGHER could be seen as a directional anti-identity constraint. It does not simply demand that the input be different than the output (in the same way that WAX does not simply require that the preceding segment be different from the following segment). Rather, it requires that the output be higher on the scale than the input. Like WAX, this type of constraint is only interpretable with reference to scales. Thus, to the extent that this family of constraints is successful in modeling existing phenomena while ruling out unattested mappings, the idea that n -ary relationships underlie phonological grammars is supported.

If the power added to the theory by DIFF was staggering, the power added by HIGHER is all the more so. Indeed, it makes it possible to replace input segments with unrelated output segments, among other things¹¹. However, it is a significant contention of this work that a grammar lacking constraints of the HIGHER type is descriptively inadequate and that all of the *types* of phenomena predicted by HIGHER do exist,

¹¹Though such replacements are not common, they do occur, especially in the tonal domain. Various types of crazy rules probably fit into this category as well, at least when they are viewed as surface phenomena.

even though it is doubtful (indeed certain) that not all of the possible *instantiations* of these types of phenomena are to be found in actual human languages.

Logically, there must be a counterpart to HIGHER, a constraint that is here called NOHIGHER. This constraint is equivalent to NOWAX and differs from HIGHER in being simpler—in consisting of the ordering relation alone and not to the negation of the ordering relation. A formal definition of HIGHER follows:

(71) NOHIGHER[S]

Assume that S is the totally ordered set $\langle \leq, T \rangle$ and that \leftarrow is a correspondence relation from an output candidate \mathbb{C} to an input \mathbb{I} ; for any $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{I}$, if $\alpha, \beta \in T$ and $\beta \leftarrow \alpha$ then $\alpha \leq \beta$.

As should be clear from the definition, the violation profile of NOHIGHER is exactly the opposite of that of HIGHER. Thus, the candidate pairs in 70 satisfy NOHIGHER while those in 69 violate it.

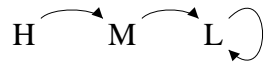
2.7 A Factorial Typology of Mapping Schemas

Having described the set of input-output constraints and the other constraints with which they interact, it is now possible to construct a factorial typology of possible input-output mappings. The following schemas are those that result from the possible rankings of SAME, DIFF, HIGHER, and ENDMOST (relative to the same scale). It will be shown that each of the types of mappings predicted to exist is attested in a natural language.

For each description, the mappings will be presented as directed pseudo-graphs where each edge (arrow) corresponds to a mapping relationship with vertex at the beginning of the arrow corresponding to the input and that at the end of the arrow corresponding to the output. For example, the graph in (72) represents a grammar where M

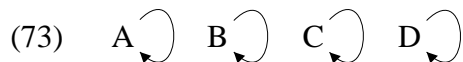
is mapped to L, H is mapped to M, and L is mapped to itself (does not alternate).

(72) **Chain shift from Eastern A-Hmao** (Wang and Wang 1986):



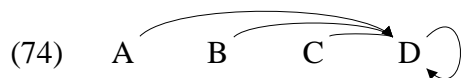
2.7.1 The Identity Mapping

If SAME is undominated, the result is the identity mapping¹². This follows from simple logic: if the most important requirement is that the input be identical to the output then any candidate pair in which the input differs from the output will be defeated by the fully faithful candidate. This state of affairs can be depicted as in (73):



2.7.2 Neutralization

If ENDMOST dominates all of the other constraints, then a neutralization to the lowest rung of the scale will always result¹³. This, like the identity mapping, can be proved without demonstration. If the most important requirement is that outputs be at the lowest point on the scale, then any output candidate which deviates from this requirement will be defeated by a candidate that resides at the lowest point on the scale. This is shown in (74):



¹²The identity mapping is generated by 6 of 24 possible rankings of the constraints under discussion (SAME, DIFF, HIGHER, and ENDMOST).

¹³The neutralization mapping is generated by 6 of 24 possible rankings.

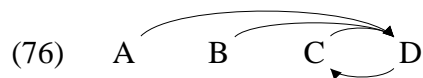
It is easy enough to find cases of input-output mappings that match this schema. A convenient example can be drawn from Dananshan Hmong (Wang 1985):



This type of mapping, however, presents an interesting type of ambiguity. It can never be ascertained, either by the language learner or the linguist, whether a given pattern of neutralization occurs across a scale of three or more increments (as presupposed by (74)) or within a two-step scale in which all of the elements but the “endpoint” are at the penultimate step on the scale (or some intermediate scenario) without independent knowledge about the structure of the scale.

2.7.3 Neutralization with “Bounce-Back”

If DIFF dominates both ENDMOST and SAME and ENDMOST dominates HIGHER then all inputs will be mapped to the end-point of the scale with the exception of inputs at the end of the scale, which will be mapped to the penultimate point on the scale¹⁴. This can be schematized as in (76):



Intuitively, this mapping results because both ENDMOST and DIFF can be satisfied as long as inputs are not at the endpoint of the scale. However, DIFF can only be satisfied by mapping inputs at the endpoint of the scale to some other point. This point will be the point which incurs the fewest violations of ENDPPOINT without actually being at the endpoint. That point is always the penultimate point on the scale. This kind of “return” to the penultimate point on the scale will be referred to here as “bounce-back.”

That the ranking

¹⁴A “neutralization with bounce-back” mapping is generated by 3 of 24 possible rankings.

(77) DIFF ≫ ENDMOST ≫ HIGHER, SAME

generates the mapping given in (76) is proved by demonstration in the tableaux in (78):

		/A/	DIFF	END	HIGHER	SAME
(78) a.	(a)	A	*!	***	*	
	(b)	B		*!*	*	*
	(c)	C		*!	*	*
	(d)	D			*	*

		/B/	DIFF	END	HIGHER	SAME
b.	(a)	A		*!***		*
	(b)	B	*!	**	*	
	(c)	C		*!	*	*
	(d)	D			*	*

		/C/	DIFF	END	HIGHER	SAME
c.	(a)	A		*!***		*
	(b)	B		*!*		*
	(c)	C	*!	*	*	
	(d)	D			*	*

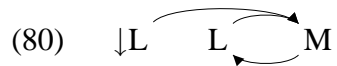
		/D/	DIFF	END	HIGHER	SAME
d.	(a)	A		**!*		*
	(b)	B		**!		*
	(c)	C		*		*
	(d)	D	*!		*	

The ranking

(79) DIFF ≫ SAME ≫ ENDMOST ≫ HIGHER

generates a mapping of the same type. Proving this, and the similar claim made above, is trivial. It is possible to satisfy DIFF for any input, so the winning candidate will never be the fully faithful candidate (thus the winning candidate will always violate SAME which is rendered ineffectual by being ranked below DIFF). Since ENDMOST dominates HIGHER, the output will always be at the endpoint of the scale except in the case where the output is at that point. In that case, the output will be the unfaithful candidate which violates ENDMOST the fewest times.

This type of mapping, strange though it may seem, is attested in a few languages. One particularly clear example is to be found in the tone system of Western A-Hmao (Johnson 1999) (see also (Wang and Wang 1986)):



In this language, the lowest tones are mapped to M, but M is mapped to the L tone rather than $\downarrow L$. That is to say, it “bounces back” to the second worst tone in terms of markedness (as evaluated by ENDMOST). It is because of cases of this type that Harmonic Ascent cannot be allowed to tie the hands of the grammatical apparatus.

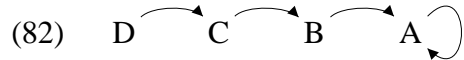
With the exception of Lai (2002) and Mortensen (2003), relatively little attention has been paid to phenomena of this type. As noted above in Section 2.6.2, Classical Optimality theory, or any version of Optimality Theory in which harmonic ascent holds, this type of mapping cannot be generated. However, it can be captured in terms of comparative markedness, but only at the expense of adding the power to generate infinite chain shifts. The scalar mechanism proposed here is able to generate the attested pattern without predicting the existence of unattested patterns such as unconditioned epenthesis.

2.7.4 Chain shift

Under the ranking

(81) HIGHER \gg SAME \gg DIFF, ENDMOST

a classical chain shift results¹⁵. The rationale is simple: outputs should be higher than their corresponding inputs. Where this is not possible (because the inputs are at the top of the relevant scale), the next best option is for outputs to be identical to the corresponding inputs. This mapping is depicted by the schema in (82):



A formal demonstration that the ranking in (81) generates this mapping is given in the tableaux below (83):

(83) a.

/D/	HIGHER	SAME	DIFF	END
(a) A		*		**!*
(b) B		*!		**
(c) C		*		*
(d) D	*!		*	

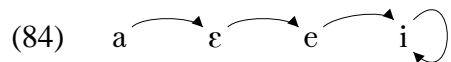
b.

/C/	HIGHER	SAME	DIFF	END
(a) A		*		***!
(b) B		*		**
(c) C	*!		*	*
(d) D	*!	*		

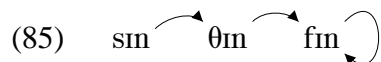
¹⁵Chain shift mappings are generated by 2 of 24 possible rankings.

	/B/	HIGHER	SAME	DIFF	END
c.	(a) A		*		***
	(b) B	*!		*	**
	(c) C	*!	*		*
	(d) D	*!	*		
	/A/	HIGHER	SAME	DIFF	END
d.	(a) A	*		*	***
	(b) B	*	*!		**
	(c) C	*!	*		*
	(d) D	*!	*		

Chain shifts of this type are widely attested. A fairly large collection of such chain shifts (as well as other examples of counter-feeding processes) has been compiled by Moreton (2004a). A perfect (four-step) example exists in Nzebi (Guthrie 1968; Clements 1991; Kirchner 1995, 1996) as shown in (84)¹⁶:



Three-step chain shifts are far more common, both in adult grammars and in the speech of children. A common chain shift in the speech of children learning English is the fricative place shift (Dinnsen and Barlow 1998):



In this shift, /s/ is realized as [θ] but /θ/ and /f/ are both realized as [f]. These chain shifts are different from many examples of counterfeeding opacity. The theoretical mechanism given here will only generate chain shifts that are type-preserving (that is,

¹⁶In the same context, Nzebi back vowels are raised in a three-step chain: ɔ → o → u.

chain shifts in which each “step” consists of the substitution of an element of some type for another element of the same type). Types of counterfeeding opacity which do not satisfy these conditions must be generated through some other mechanism. As will be argued below, there are reasons for distinguishing these two types of processes (true—type preserving—chain shifts and epiphenomenal counterfeeding).

2.7.5 Ring (circular chain shift)

The input-output constraints described here are able to generate circular chain shifts (“rings”) under seven different rankings, all of which meet the following condition:

$$(86) \quad \text{HIGHER} \gg \text{ENDMOST} \wedge (\text{ENDMOST} \gg \text{SAME} \vee \text{DIFF} \gg \text{SAME})$$

That is to say HIGHER dominates ENDMOST and either ENDMOST or DIFF dominates SAME. This ranking can be expanded into the following three (tableau friendly) rankings:

- (87) a. HIGHER \gg ENDMOST \gg DIFF, SAME
 b. HIGHER, DIFF \gg ENDMOST, SAME
 c. DIFF \gg SAME \gg HIGHER \gg ENDMOST

Under all of these rankings¹⁷, HIGHER is able to “drive” the shift up the scale. Also, in each of these rankings, the “rounding of the horn,” that is, the mapping from the top of the scale to the bottom, is driven by ENDMOST. This results in a mapping like that in example (88).



¹⁷The “ring” mapping is generated by 7 of 24 possible rankings.

Proving that the rankings in (87) generate the mapping in (88) requires us to solve each of the cases individually.

The first case is the ranking

(89) HIGHER \gg ENDMOST \gg DIFF, SAME

which, of course, subsumes two strict rankings. The tableaux in (90) prove that these rankings generate the mapping given in (88), that is, that they generate a ring mapping.

		/D/	HIGHER	END	DIFF	SAME
(90) a.	(a) A			**!*		*
	(b) B			**!		*
	☞ (c) C			*		*
	(d) D		*!		*	
		/C/	HIGHER	END	DIFF	SAME
b.	(a) A			****!		*
	☞ (b) B			**		*
	(c) C		*!	*	*	
	(d) D		*!			*
		/B/	HIGHER	END	DIFF	SAME
c.	☞ (a) A			****		*
	(b) B		*!	**	*	
	(c) C		*!	*		*
	(d) D		*!			*

	/A/	HIGHER	END	DIFF	SAME
d.	(a) A	*	*!***	*	
	(b) B	*!	**		*
	(c) C	*!	*		*
	☞ (d) D	*			*

Likewise, the tableaux in (92) prove that the ranking (or set of four strict rankings)

(91) HIGHER, DIFF \gg ENDMOST, SAME

produces a circular chain shift.

	/D/	DIFF	HIGHER	END	SAME
(92) a.	(a) A			***!	*
	(b) B			**	*!
	☞ (c) C			*	*
	(d) D	*!	*		

	/C/	DIFF	HIGHER	END	SAME
b.	(a) A			***	*!
	☞ (b) B			**	*
	(c) C	*!	*	*	
	(d) D		*!		*

	/B/	DIFF	HIGHER	END	SAME
c.	☞ (a) A			***	*
	(b) B	*!	*	**	
	(c) C		*!	*	*
	(d) D		*!		*

	/A/	DIFF	HIGHER	END	SAME
d.	(a) A	*	*!	***	
	(b) B		*!	**	*
	(c) C		*!	*	*
	(d) D		*		*

The final case, where the ranking is

(93) DIFF \gg SAME \gg HIGHER \gg ENDMOST

is shown to produce a chain by the tableaux in (94):

	/D/	DIFF	SAME	HIGHER	END
(94) a.	(a) A		*		**!*
	(b) B		*!		**
	(c) C		*		*
	(d) D	*!		*	

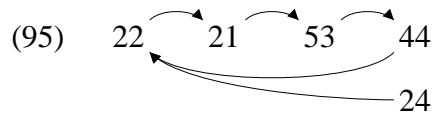
	/C/	DIFF	SAME	HIGHER	END
b.	(a) A		*		***!
	(b) B		*		**
	(c) C	*!		*	*
	(d) D		*!	*	

	/B/	DIFF	SAME	HIGHER	END
c.	(a) A		*		***
	(b) B	*!		*	**
	(c) C		*	*!	*
	(d) D		*!	*	

	/A/	DIFF	SAME	HIGHER	END
d.	(a) A	*!		*	***
	(b) B		*	*	*!*
	(c) C		*!	*	*
	(d) D		*	*	

It is notable that so many rankings (seven of them, in fact) generate this mapping, especially in light of the fact that this type of mapping is so uncommon. However, there is one very well-known example of a circular chain shift, namely the circle shifts of various dialects of Southern Min (including Taiwanese and Xiamen).

The Taiwanese tone circle is depicted in (95), with Chao tone numbers representing the values of the tones¹⁸:



The other known cases of circular chain shifts having more than two points are all tone shifts and are all found in Southern Min languages (see Ballard 1988). It is notable, however, that this analytically-challenging phenomenon falls out as a natural prediction from the theoretical mechanism that has been required to account for other, more robustly attested, phenomena.

2.7.6 Comparison of Typological Predictions

We have seen that a Structural Optimality grammar containing the constraints SAME, DIFF, HIGHER, and ENDMOST can generate five different types of mappings: iden-

¹⁸In this system (Chao 1930), “5” represents the highest pitch and “1” the lowest pitch, with level tones represented by a sequence of two identical numbers and contour tones represented by a sequence of two or more non-identical numbers.

tivity mappings, endpoint neutralizations, neutralizations with bounce-back, finite chain shifts, and circular chain shifts. All of these mappings, we have seen, correspond to patterns in actual languages, and—taken broadly—all mappings between underlying inventories and surface inventories that are found in natural languages consist of patterns of these types.

Classical optimality theory, without constraint conjunction or similar mechanisms, predicts the existence of neutralization mappings and identity mappings only. Once local constraint conjunction is added to the mix, an Optimality Theory constraint set can generate identity mappings, neutralizations, and chain shifts (as well as being able to capture a variety of different types of phonological opacity). Harmonic Ascent holds for such grammars, so they can never generate bounce-back effects, rings, or infinite chain shifts.

In contrast, Harmonic Ascent does not hold for OT grammars including comparative markedness constraints. Among such grammars are those that generate identity mappings, neutralizations, chain shifts, and bounce-back neutralizations. Some grammars of this type also generate infinite chain shifts, which appear to be unattested in natural languages. These grammars do not appear able to generate the type of circular chain shifts where the loop contains more than two points (designated as “rings” here).

In contrast, the constraint set described here generates all and only the attested mapping patterns. In so doing, it provides a partial solution to the opacity problem in Optimality Theory, accounting beautifully for those types of counter-feeding opacity that are type-preserving (that is to say, classical chain shifts). It does, by itself, provide an explanation for other types of opacity (non-type preserving counterfeeding, counterbleeding, and derived environment effects).

Chapter 3

Chain Shifts and Scales

It is not a field of a few acres of ground,
but a cause, that we are defending, and
whether we defeat the enemy in one
battle, or by degrees, the consequences
will be the same.

“The Crisis”
Thomas Paine

3.1 Chain Shifts: Problems and Prospects

Chain shifts have always presented an interesting problem in phonological theory. On the one hand, they are easy to formalize: on the other hand, formalizations of chain-shifts have not tended to capture the insight that certain chain shifts function as a single process. In theories with ordered rules, chain shifts are easily modeled with a sequence of two or more rules. Take a case where underlying /a/ is mapped to surface [e] but underlying /e/ is mapped to surface [i]. Using ordered rules, we would reduce this chain to two processes:

- (96) a. /e/ → [i]
 b. /a/ → [e]

It is notable that the only connection between these rules is the epiphenomenal interaction between them. They are no more connected than the rules that would be used to model a rather different type of counter-feeding opacity. Take the famous example of

/katab/ ‘to write’ and /badw/ ‘Bedouin’ in Bedouin Arabic. In this language there is a process that raises underlying /a/ to [i] in open syllables and another process of glide vocalization. The process of glide vocalization renders vowel raising “surface untrue” since /badw/ features an /a/ in an open syllable:

(97)	/badw/	/katab/	
	badw	ka.tab	Syllabification
	badw	ki.tab	Raising of /a/ in open syllables
	badu	—	Glide vocalization
	[ba.du]	[ki.tab]	

According to McCarthy’s (1999; 2002) representation of this case, we could expect /badw/ to surface as [badu] but /badu/ to surface as [bidu].

Our intuition is that there is a difference between these two types of processes—scalar (chain-shifting) and non-scalar. One appears to arise from a single source (vowel reduction, for example, assimilation between vowels, or breathy phonation—causes that will be investigated later). The other results from two separate processes that, like the rules that could be used to model them, are related only by the ordering interaction between them.

This same problem is shared by certain contemporary approaches to counter-feeding couched in Optimality Theory. In Sympathy Theory and Comparative Markedness (McCarthy 2002), opacity is modeled by introducing an intermediate level of representation (the sympathy candidate and the fully-faithful candidate, respectively) parallel to the output. Like the rule-ordered approach to chain-shifts, this kind of model divides chain-shifts into separate processes. There are three notable exceptions to this trend: Gnanadesikan’s (1997) use of ternary scales to model chain-shifts, Kirchner’s (1995; 1996) distantal faithfulness model and Łubowicz’s (2003) contrast-preservation theory of chain shifts, which captures a different (and interesting) insight about the nature of

chain-shifting opacity.

Just as there are two types of counter-feeding opacity in synchronic chain shifts, there are two classes of chain-shifts in diachronic phonology¹. One type consists of two sound changes whose mutual relationship is purely coincidental. The second type consists of two sound changes that were triggered by the same conditioning factor. A central argument of this study is that true synchronic chain shifts usually arise from diachronic chain shifts of exactly this type—that they represent a single process with a single cause, rather than a sequence of processes.

Synchronically, Chain Shifts result when HIGHER dominates SAME, which in turn dominates DIFF and ENDMOST. HIGHER demands that the output be higher along the relevant scale than the input. It is the force that motivates movement towards one end of the scale. If the input is at the high end of the scale already, the grammar must settle for the next best option—the case in which the input and the output are at identical points on the scale. The steps up the scale will always be minimal since minimal jumps are better relative to ENDMOST but have identical violation profiles relative to HIGHER and SAME. DIFF ends up being irrelevant because it is dominated by SAME.

This can be made more clear with a concrete example. Assume the following scale

$$(98) \quad H=\{i\}_2 > \{e\}_1 > \{a\}_0$$

The input-output pair (/a/, [i]) satisfies HIGHER, violates SAME, and incurs two violations of ENDMOST. This candidate is less optimal than the pair (/a/, [e]), since this pair also satisfies HIGHER while only violating ENDMOST once and does not differ from (/a/, [i]) in relative to SAME:

¹More accurately, there are two types of *conditioned* chain shifts in diachronic phonology.

	a	HIGHER(<i>H</i>)	SAME(<i>H</i>)	END(<i>H</i>)	DIFF(<i>H</i>)
(99)	(a) a	*!			*
	(b) e		*	*	
	(c) i		*	**!	

For the input /e/, only output [a] can satisfy HIGHER, so it is selected as the output:

	e	HIGHER(<i>H</i>)	SAME(<i>H</i>)	END(<i>H</i>)	DIFF(<i>H</i>)
(100)	(a) a	*!	*		
	(b) e	*!	*	*	
	(c) i		*	**	

For the input /i/, no output can satisfy HIGHER, so the grammar does the next best thing: it selects the output candidate [i], which satisfies the next-highest ranked constraint:

	i	HIGHER(<i>H</i>)	SAME(<i>H</i>)	END(<i>H</i>)	DIFF(<i>H</i>)
(101)	(a) a	*	*!		
	(b) e	*!	*	*	
	(c) i	*		**	

Any simple chain shift can be generated by this ranking schema.

Having walked through an overview of both the motivations for the treating chain shifts as scalar phenomena and the formal structure of the analyses, it is now essential that we look at several actual chain shifts.

3.2 Western Hmongic Tonal Chain Shifts

Tonal chain shifts are quite common in the languages of China and Southeast Asia, specifically among languages with tone sandhi of the “Chinese” type². For example, a

²Though it is probably not helpful to divide tonal processes into absolute categories, it seems that some tonal systems do have different properties than others. Many (though not all) Sinitic languages have a type of tone sandhi in which one tone is replaced by another in some context. This type of

number of examples from Chinese languages are described by Ballard (1988). Here I will argue that a significant number of these chain shifts originated from a consistent conditioning factor perturbing multiple points along the pitch grade, that is, that they result from the same types of factors as true chain shifts generally. These shifts, I will demonstrate, begin as sets of phonetically coherent alternations but may develop (through subsequent changes) into phonetically arbitrary processes.

A particularly striking illustration of this point can be made on the basis of a tone sandhi chain found in Western Hmongic languages. The original circumstances that lead to the emergence of this chain are best preserved in Dananshan, but the same system exists, in slightly obscured form, in most of the other language varieties of the Western Hmongic group³. Its evolution may be traced from the Proto-Western Hmongic form (preserved in Dananshan and, to a lesser extent, A-Hmao) to the complicated and arbitrary looking shifts found in the Mashan⁴ dialects of Xinzhai and Shuijingping.

3.2.1 Dananshan Hmong

Dananshan is a dialect of Hmong spoken in Dananshan village, Guizhou, China. Hmong belongs to the group of languages that are called *Hmongic* by Western linguistics and

process lies in contrast to tonal processes found in certain African, Mesoamerican, Tibeto-Burman, and even Sinitic, languages where tonal alternations can be easily characterized by processes like spreading, contour simplification, downstep, and so forth.

³For a delineation of the membership of this group, see Johnson (2002). For earlier comparative observations relating to this chain shift, see Downer (1967) and Ratliff (1992b,a)

⁴The Mashan dialects form one subdivision of the Western Hmongic family. They are genetically closer to Far Western Hmongic languages (such as Dananshan, Hmong Daw, Mong Leng, and A-Hmao) than the Guiyang dialects but nevertheless lie outside the Far Western Hmongic group, as can be demonstrated by developments in the tone sandhi system that are shared by all Far Western Hmongic languages but not by the Mashan dialects. For primary data, see Niederer (1998); for a subgrouping based on an analysis of these data, see Mortensen (2005).

are known as *Miao* by linguists in China. Like many other Hmong-Mien (Miao-Yao) languages, Dananshan has a large tonal inventory and an interesting set of tonal alternations.

The complete inventory of contrasting tonal melodies on Dananshan monosyllables is given in (102):

- (102) a. Level: ↑H, H, M, L
 b. Contour: HM, MH, ML, LM

The tone sandhi system of Dananshan consists entirely of tone changes that occur following a falling tone—either HM or ML. Certain of these changes have a very long pedigree in Western Hmongic. These are reflected, in Dananshan, as the neutralization of L, ML, LM, and MH as LM. The changes which will concern us most, though, have developed fairly recently (demonstrated by the fact that they only occur in one genetic subgroup within Far-Western Hmongic, the Chuanqiandian dialect group). They form a lowering chain in which ↑H (super high) becomes H (high) and H becomes M (mid) following a falling contour tone (Wang 1985; Niederer 1998).

Concretely, ↑H is lowered to H after both HM and ML (the only falling contours in the language), as shown in example (103):

- (103) a. ntou^{HM} ‘cloth’ sa^{↑H} ‘blue’ ntou^{HM}sa^H ‘blue cloth’
 b. t̚si^{HM} ‘five’ teu^{↑H} ‘peck’ t̚si^{HM}teu^H ‘five pecks’
 c. ku^{ML} ‘trench’ t̚se^{↑H} ‘house’ ku^{ML}t̚se^H ‘sewer’
 d. nqai^{ML} ‘flesh’ tle^{↑H} ‘dog’ nqai^{ML}tle^H ‘dog flesh’

In the same context (following HM and ML) H is lowered to M:

- (104) a. au^{HM} ‘two’ pua^H ‘hundred’ au^{HM}pua^M ‘two hundred’
 b. plou^{HM} ‘hair’ npua^H ‘pig’ plou^{HM}npua^M ‘pig hair’
 c. nploŋ^{ML} ‘leaf’ ntoŋ^H ‘tree’ nploŋ^{ML}ntoŋ^M ‘tree leaf’
 d. na^{ML} ‘year’ ɕoŋ^H ‘year’ na^{ML}ɕoŋ^M ‘year’

This type of tone change occurs both within polysyllabic words (most of which are compounds) and between words within the prosodic word. The most common extra-lexical domain for tone sandhi in Dananshan and related languages is between a numeral and a noun classifier. It is important to note that these two elements do not form a morphological or syntactic constituent. While the process is circumscribed by a particular context, this context is not strictly lexical⁵.

3.2.1.1 Origin of the Chain Shift

It is not difficult to image how a chain shift like that seen in Dananshan could originate. The falling tones (which may have originally been realized as ML and ML) had a depressing effect upon the tone of a following syllable as a result of tonal coarticulation. This seems to accord generally with the two most robust facts about tonal articulation: first, that tonal coarticulation (at least in Chinese languages) tends to be perseveratory rather than anticipatory, and second, that perseveratory coarticulation tends to be assimilatory while anticipatory effects tend to be dissimilatory (Xu 1997). The glottal gesture necessary to produce a falling tone would seem to militate against gestures needed to raise the pitch of the following syllable to H (high) or ↑H (super high). This automatic lowering of tones following falling contours would thus be most noticeable on the higher tones (e.g. H and ↑H). This automatic effect could then be phonologized. Phonologization of this pitch perturbation was probably aided by the fact that there were already a set of (much older) tone sandhi alternations conditioned by the same two tones, as described above (Wang 1985; Mortensen 2005). If this is the correct approach to this problem, it follows that the emergence of these alternations was the result of a conspiracy between phonetic factors and grammatical factors. It would be an important example of the way grammatical knowledge can bias the inferences that

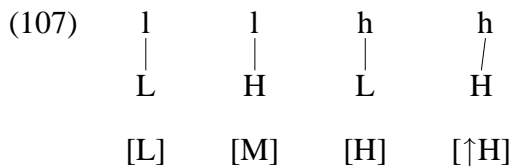
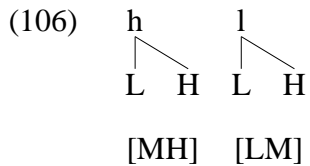
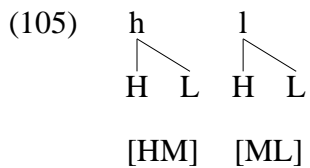
⁵This could be seen, of course, as “precompiled phrasal phonology,” as argued for by Hayes (1990).

speakers take from their sensory experience of speech. On the one hand, there seems to be a clear phonetic motivation for the lowering effect. On the other hand, there were already tonal alternations in the same environment, that did not involve the same pattern of change. The fact that the new, phonetically grounded, alternations emerged in exactly the same environment as the existing tone sandhi alternations is too striking to be coincidental.

This chain-shifting pattern seems to have originated as a kind of uniform perturbation across a phonetic grade. Tone is a naturally gradient dimension—pitch is a continuous rather than a discrete parameter. Other factors that are sometimes important in distinguishing tonal categories, such as voice quality and duration, are also continuously variable. The depressing effect of a preceding contour applied to elements at various points within the pitch grade, so that both the \uparrow H tone and the H tone were subject to a similar effect. This account of the Western Hmong tonal chain shift is superior to an account that relies solely upon conventional feature geometries, as we will now demonstrate.

3.2.1.2 Problems with an Autosegmental Account

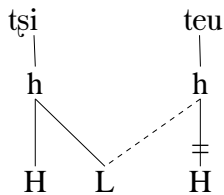
Dananshan Hmong has a tonal inventory that seems to match that predicted by geometric theories of tone like those of Yip (1980) and Bao (1999) exactly. In this language, there are two rising tones (represented as in 106), two falling tones (as in 105) and four level tones (as in 107). This is exactly the tonal inventory that would result if one accepted Bao's (1999) assertion that the sequence of tone features on a TBU (tone bearing unit) is maximally binary, that there is one binary feature for tone melodies and one binary feature for register. This would result in a set of eight tones: two falling tones (105), two rising tones (106), and four level tones (107). This matches the tonal inventory of Dananshan exactly:



However, as shall be seen, these tonal representations are not helpful in accounting for certain tonal phenomena within this language. This is not to say that they are useless for the language generally, or even that they are wrong, but only that some other representational device is needed. This device is a scale.

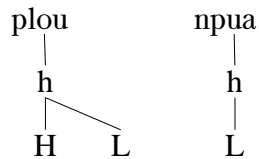
Assuming these representations, it is not terribly difficult to model the first step in the scale. As shown in (108), we could view the change from ↑H to H as the result of the spreading of a L from the preceding syllable. This spreading would be accompanied by the delinking of the H on the second syllable. Formulating a rule such that this spreading occurs just in case the L tone is preceded on the same TBU by a H may be somewhat awkward, but it is by no means impossible.

(108) $HM + \uparrow H \rightarrow HM + H$



The problem comes when we look at the next step in the chain. In that case a low register H must become a high register L. This does not seem possible to model using autosegmental spreading and the chosen system of representations:

(109) $HM + H \rightarrow HM + M$



There is not a source for a l register feature, nor a readily available source for a H tone feature for the second TBU. While it would be possible to simply stipulate changes in the values of the features, this would be unsatisfying. It would not tell us anything about the process, and it would disguise the fact that this is a unified process with a single substantive correlate, namely, an incremental decrease in pitch. For better or for worse, such an account would have no connection to the diachronic scenario we have proposed for the development of this pattern.

What would a better analysis of this pattern look like? It would almost certainly have to involve some hierarchical relationship (that I will call *T*) among $\uparrow H$, H, and M such that $\uparrow H$ is most *T*, H is less *T* than $\uparrow H$, but more *T* than M, and M is least *T* of all. This could be implemented a number of different ways. For example, we could propose some tone feature [raise] that allows recursive association (so that an instance of this feature can be associated in feature geometry with another instance of the same feature). This kind of solution has been discussed by Clements (1991) for Nzebi vowel raising (though Clement rejects such an analysis in favor of a geometric model allowing multiple [open] features). What such a solution would do is express—in an indirect way—a scalar relationship among the tones $\uparrow H$, H, and M. The solution adopted here is to state the scale directly. To wit, there is a scale *T*, a subset of which

is the relation:

$$(110) \quad \{M\}_2 > \{H\}_1 > \{\uparrow H\}_0$$

The structure of the full scale will not concern us here. What will concern us is the relationship between the motivations for the pattern we observe and the formal grammar we will construct from the constraint families explained above. The force-depressing level tones after a falling tone can be formalized as a HIGHER constraint which we will call HIGHER[*T*]. It states that output tones are higher along the scale *T* than their corresponding inputs. This competes with SAME[*T*], corresponding to the force that maintains identity between the input and output tones and ENDMOST[*T*], which corresponds to the force antagonistic to the lowering effect of HIGHER[*T*]⁶ that prevents a complete neutralization of all of the tones on the scale to M⁷. It counts one violation for each stem between then end of the scale (where M resides) and any tone in the output. Thus, an output [H] would violate ENDMOST[*T*] once and an output [\uparrow H] would violate ENDMOST[*T*] twice.

The ranking of these constraints follows the same simple schema we have already discussed. HIGHER[*T*] is undominated: if possible, outputs should be higher than their corresponding inputs. SAME[*T*] is ranked just below HIGHER[*T*], meaning that—in the event that it is not possible to go any farther up the scale—the next best move is not to change at all. The ranking of ENDMOST[*T*] is irrelevant as long as it is dominated by both HIGHER[*T*] and SAME[*T*]. If ENDMOST[*T*] was to dominate SAME[*T*], a ring

⁶The irony of employing a constraint called HIGHER to produce a lowering effect does not escape the author. The scale could easily be reversed, this constraint could be renamed LOWER, and ENDMOST could be changed to TOPMOST, yielding exactly the same results, but in the interest of terminological consistency, the current scheme has been employed.

⁷In actuality, it is not necessarily true that a grammar not containing ENDMOST[*T*] would map both \uparrow H and H to M. H will always be mapped to M, but the winning candidate for the output corresponding to \uparrow H would be indeterminate unless some other low-ranked constraint were invoked. The presence of ENDMOST[*T*] resolves this indeterminacy

would result: when the scale could be ascended no further, the next best option would be to go to the bottom end of the scale. The same result would be had if $\text{DIFF}[T]$ was to dominate $\text{SAME}[T]$. The results of these rankings are shown in Tableaux (111-113):

	$\uparrow\text{H}$	$\text{HIGHER}(T)$	$\text{SAME}(T)$	$\text{END}(T)$	$\text{DIFF}(T)$
(111)	(a) $\uparrow\text{H}$	*!			*
	(b) H		*	*	
	(c) M		*	**!	
	H	$\text{HIGHER}(T)$	$\text{SAME}(T)$	$\text{END}(T)$	$\text{DIFF}(T)$
(112)	(a) $\uparrow\text{H}$	*!			
	(b) H	*!		*	*
	(c) M			**	
	M	$\text{HIGHER}(T)$	$\text{SAME}(T)$	$\text{END}(T)$	$\text{DIFF}(T)$
(113)	(a) $\uparrow\text{H}$	*	*!		*
	(b) H	*!		*	
	(c) M	*		**	

This proposal compares favorably with earlier approaches to chain shifts in Optimality Theory advanced by Kirchner (1995, 1996), Gnanadesikan (1997), and Dinnsen and Barlow (1998). It is true that this analysis shares something with each of these earlier proposals. Like Gnanadesikan's and Dinnsen and Barlow's⁸ proposals, the current analysis depends upon the existence of scalar representations or relations. As in Kirchner's (1995) proposal, the chain shift here is claimed to be driven by a single constraint

⁸The analysis alluded to here employs fully-specified adult representations. They reject this proposal in favor of a different analysis in which uses underspecified representations mirroring, they argue, the representations of children producing these forms (Dinnsen and Barlow 1998:84ff). Part of their reason for rejecting the first analysis is that they believed the scale that they posited was an ad hoc mechanism with no motivation. This is not a problem in the current theory.

(his RAISING constraints). However, all of these earlier approaches share a liability in the form of “distantial faithfulness” (Dinnsen and Barlow 1998). In each of these analyses, there is a faithfulness constraint that penalizes outputs that deviate from the input beyond a specified threshold. Both Gnanadesikan and Dinnsen and Barlow formulate this constraint in terms of scales, while Kirchner sees distantial faithfulness as a constraint over a continuously variable space. Kirchner, in fact, composes his equivalent of distantial faithfulness out of two atomic constraints by means of constraint conjunction. In either case, a constraint of this kind allows for chain shifts by allowing outputs to be only one increment from the input. If the constraint(s) driving the chain shift was undominated, the result would be neutralization. If it was dominated by an ordinary faithfulness constraint, the identity mapping would prevail. A distantial faithfulness constraint can be satisfied while allowing the output to improve relative to markedness.

Both Gnanadesikan (1997) and Dinnsen and Barlow (1998) model counterfeeding opacity using a constraint family (IDENT-ADJ and DISTFAITH respectively) that has no real motivation outside the domain of chain shifts⁹. These constitute a type of faithfulness constraint that does not distinguish between outputs that are identical to the input and those that are adjacent to the input on some scale. Gnanadesikan:78 defines IDENT-ADJ as follows:

(114) IDENT-ADJ[X scale]

Given an input segment α and its correspondent output segment β , then α and β must have related values on scale x, where the defined relations are identity and adjacency. (In other words, the output may not have moved more than one step on the scale.)

⁹Gnanadesikan (1997) also employs IDENT-ADJ in her analysis of “attraction” effects. These, however, can be modeled without this type of constraint.

This constraint type is useful for only a limited number of phenomena, primarily chain shifting phenomena. In contrast to these constraints, the constraint families employed here are motivated by a whole host of phenomena. A constraint of this type seems necessary in modeling ring shifts (circular chain shifts with more than two points). In its syntagmatic formulation (namely WAX), a constraint of this type is needed to account for phonologically driven ordering effects in coordinate compounds. SAME is simply an IDENT constraint in scalar guise. As will be seen below, it is necessary to the account of “register-preserving” effects in tone sandhi systems, among other things. In its syntagmatic incarnation, it is an AGREE constraint of the type needed to account for certain harmony phenomena. As for ENDMOST, it is simply a scalar markedness constraint and is needed to account for both neutralization effects and ring shifts. DIFF, while not essential to the account of chain shifts, is also motivated by bounce-back effects (along the input-output dimension) and long-distance anti-harmony effects (syntagmatically). In other words, the account given here avoids introducing any ad hoc relations among corresponding elements on the order of distancial faithfulness.

Kirchner’s model is far more economical. Using PARSE constraints which make direct reference to the duration of vowels¹⁰, which he suggests are independently motivated, he composes what is essentially a distancial faithfulness constraint by means of constraint conjunction. Constraint conjunction, likewise, is suggested by Kirchner to be independently necessary to account, among other things, for stress clash avoidance. However, constraint conjunction itself, is problematic in that it predicts the existence of a host of highly unusual phenomena. Research by Moreton and Smolensky (2002) shows that constraint conjunction makes some correct yet interesting predictions about possible chain shifts. However, there is still good reason to believe that local constraint conjunction is too powerful a mechanism without some additional constraints on its

¹⁰Kirchner (1996, 1995) attempts to account for two vowel-raising chain shifts, one in Basque and the other in Nzɛbi.

operation. A more serious problem with Kirchner’s approach is its extreme phonetic literalism. It rests, to a good degree, on access to a phonetic dimension (in his case studies, duration) and thus does not adapt transparently to chain shifts that lack a corresponding phonetic dimension.

We will now see that there are chain shifts of exactly this type, which nevertheless have the same logical structure as the natural chain shift in Dananshan. In other Western Hmongic languages, the chain shift that is so perfectly preserved in Dananshan has changed in form. In some cases (like A-Hmao), the general shape of the chain shift has remained the same—only the conditioning environment has been altered. However, in the Mashan dialects, a more radical set of changes has taken place.

3.2.2 Eastern A-Hmao

Eastern A-Hmao, which like Dananshan is a member of the Far-Western Hmongic group, has a chain shift that is directly cognate to the one found in Dananshan. However, rather than being a shift of the form

$$(115) \quad \uparrow H \rightarrow H \rightarrow M$$

it appears instead as

$$(116) \quad H \rightarrow M \rightarrow L$$

Relative to the Dananshan chain shift—almost certainly the diachronically prior pattern¹¹—the tonal values in the A-Hmao chain are more widely spread (across the whole

¹¹The reasons for this inference are as follows: Both Dananshan and Western A-Hmao (for which, see Johnson (1999)) have the “ $\uparrow H \rightarrow H \rightarrow M$ ” shift, despite the fact that they belong to different branches of Far Western Hmongic. Furthermore, the register split in Western Hmongic would have produced, at its inception, a sequence of relatively high tones and another of relatively low tones. All of these tones in the chain shift belong to the high register. As such, a tone chain which consists solely of relatively high tones is very likely to be closer to the original chain shift.

tone space, rather than just across the upper register). This kind of change is unsurprising. Tonal values in East Asian tone languages tend to wander quite freely, and small changes in the relative pitch of tonal categories are to be expected. Note, however, that this change in tonal values has not altered the relationship that held between the tones in the chain.

- (117) a. tu^H ‘son’ ki^H ‘grandchild’ tu^Hki^M ‘descendants’
 b. mau^H ‘Miao’ sa^H ‘Han’ mau^Hsa^M ‘citizenry’
 c. ti^H ‘land’ tɕhœy^M ‘place’ ti^Htɕhœy^L ‘location’
 d. qu^H ‘old’ tʂho^M ‘clothing’ qu^Htʂho^L ‘old clothing’

It is notable, however, that the conditioning environment for the sandhi alternations seems to have changed. In the examples in (117), the tone of the syllable preceding the alternating tone is always H¹², rather than the falling tones we saw in Dananshan. Additionally, the MH tone triggers the same sandhi alternations as H:

- (118) a. ɣfiu^{MH} ‘lusheng’ ʂu^H ‘sound’ ɣfiu^{MH}ʂu^M ‘lusheng sound’
 b. lfi^{MH} ‘long time’ nti^H ‘long’ lfi^{MH}nti^M ‘for a long time’
 c. ɲfiie^{MH} ‘year’ ɕau^M ‘year’ ɲfiie^{MH}ɕau^L ‘age’
 d. dzfiie^{MH} ‘animal’ mpa^M ‘pig’ dzfiie^{MH}mpa^L ‘beast of burden’

This fact appears unsettling until it is revealed that the H that conditions sandhi is cognate to the Dananshan HM tone and the A-Hmao MH tone is cognate to the Dananshan ML tone. In other words, the substance of the tones conditioning the sandhi alternations has changed, but their behavior relative to the other tones has not. If we were to give this chain shift an analysis unrelated to the analysis proposed for Dananshan, we would miss the generalization that we are observing essentially the same process.

¹²There are two kinds of Highs in Eastern A-Hmao: one which conditions the tone sandhi alternations described here but does not lower to M in sandhi context and one that does not condition these alternations but does lower to M in sandhi context. These are cognate to Dananshan HM and ↑H respectively.

What will be seen next is that the tones in the chain itself may change quite markedly without changing the structure of the chain.

3.2.3 Mashan Dialects

In the Mashan dialects of Western Hmongic, the relationship between the steps in the chain has been altered considerably. In Xinzhai Hmong (Xian 1990; Niederer 1998), for example, the chain now appears as¹³:

(119) MH →LM →ML

in the context following M, LM, and H. That is to say—in this environment—MH goes to LM:

- (120) a. zei^M ‘honey’ muu^{MH} ‘bee’ $zei^M muu^{LM}$ ‘bee honey’
 b. no^{LM} ‘day’ no^{MH} ‘this’ $no^{LM} no^{LM}$ ‘today’
 c. $\eta t\epsilon a^H$ ‘post’ pla^{MH} ‘house’ $\eta t\epsilon a^H pla^{LM}$ ‘pillar’

while LM goes to ML:

- (121) a. $t\epsilon u^M$ ‘boy’ zou^{LM} ‘young’ $t\epsilon u^M zou^{ML}$ ‘young boy’
 b. zeu^{LM} ‘heart’ zau^{LM} ‘good’ $zeu^{LM} zau^{ML}$ ‘conscience’
 c. $so\eta^H$ ‘fat’ mpo^{LM} ‘pig’ $so\eta^H mpo^{ML}$ ‘lard’

Here, the lowering tendency is still in evidence for the first step on the chain shift. However, the second step looks instead like a “metathesis” of the contour. This situation undoubtedly results from the free movement of “tones” (historical tonal categories) within a multidimensional tone-space.

Data from Shuijingping Hmong illustrate a much more radical development. In the first of two chain shifts, HM becomes $\uparrow H$ in sandhi context, as shown in (122):

¹³To be precise, there are actually two chain shifts in place of one in all of the Mashan dialects. This matter will be taken up in more detail in the discussion of Shuijingping.

- (122) a. ka^M ‘medicine’ tɕə^{HM} ‘liquor’ ka^Mtɕə^{↑H} ‘brewer’s yeast’
 b. zɕei^M ‘honey’ moŋ^{HM} ‘bee’ zɕei^Mmoŋ^{↑H} ‘bee honey’
 c. na^L ‘bag’ ntə^{HM} ‘book’ na^Lntə^{↑H} ‘book bag’
 d. noŋ^L ‘day’ na^{HM} ‘this’ noŋ^Lna^{↑H} ‘today’
 e. ntɕa^{HL} ‘post’ pra^{HM} ‘house’ ntɕa^{HL}pa^{↑H} ‘pillar’
 f. tɕoŋ^{HL} ‘guide’ ka^{HM} ‘road’ tɕoŋ^{HL}ka^{↑H} ‘show way’

↑H, on the other hand, becomes H (123):

- (123) a. he^M ‘chicken’ he^{↑H} ‘crow’ he^Mhe^H ‘cock’s crow’
 b. ŋkɔ^M ‘yolk’ ha^{↑H} ‘egg’ ŋkɔ^Mha^H ‘egg yolk’
 c. lɔ^L ‘leg’ zɹ^{↑H} ‘little’ lɔ^Lzɹ^H ‘calf’
 d. ti^L ‘RECIP’ he^{↑H} ‘curse’ ti^Lhe^H ‘quarrel’
 e. poŋ^{HL} ‘grotto’ pra^{↑H} ‘cliff’ poŋ^{HL}pra^H ‘grotto’
 f. tɕain^{HL} ‘complete’ paŋ^{↑H} ‘air’ tɕain^{HL}paŋ^H ‘catch cold’

Here, the members of the chain have themselves undergone a kind of circular chain shift in their history. *↑H has become a falling tone HM (like the old tone A1), *H has moved into its place as ↑H and M too has been raised to H. *HM (historical A1) ends up as M, completing the chain. This migration of tonal categories around the tonal space has greatly altered the substance of the tones involved in this process, but the relationship between these categories has been preserved. A chain shift which was phonetically natural and well-motivated at its birth has become phonetically opaque.

In the second of the chains, M becomes the rising-falling tone LML (124):

- (124) a. kua^M ‘bug’ ntsɔ^{LML} ‘louse’ kua^Mntsɔ^{MH} ‘head louse’
 b. hei^M ‘dig’ hoŋ^{LML} ‘hole’ hei^Mhoŋ^{MH} ‘dig a hole’
 c. na^L ‘ear’ soŋ^{LML} ‘rice’ na^Lsoŋ^{MH} ‘ear of rice’
 d. mu^L ‘NEG’ laŋ^{LML} ‘grow’ mu^Llaŋ^{MH} ‘not grow’
 e. nɔ^{HL} ‘eat’ hua^{LML} ‘full’ nɔ^{HL}hua^{MH} ‘eat until full’

Oddly, LML becomes MH (125):

- (125) a. ʔei^M ‘one’ li^{MH} ‘month’ ʔei^Mli^{LM} ‘one month’
 b. ʈua^M ‘CLF’ lu^{MH} ‘the hoe’ ʈua^Mlu^{LM} ‘hoe’
 c. mu^L ‘NEG’ ɳaŋ^{MH} ‘believe’ mu^Lɳaŋ^{LM} ‘not believe’
 d. tɕa^{HL} ‘nine’ mo^{LH} ‘night’ tɕa^{HL}mo^{LM} ‘nine nights’

Thus, the two chain shifts are as follows:

- (126) a. HM → ↑H → H
 b. LML → MH → ML

In the Mashan dialects, there is a secondary tonal split affecting four of the tones, the “high” tonal register that includes all of the tones that participate in the lowering chain shift discussed here and above¹⁴. In this split, the tones of syllables with aspirated onsets¹⁵ were lowered, and aspiration was subsequently lost. The result of this split, the chain shift was “split” as well. In other words, there were then two chain shifts reflecting the historical pattern rather than just one.

The second step in both chains preserves the lowering tendency we observed in Dananshan and A-Hmao. This contrasts with Xinzhai, where the lowering tendency is manifest in the first step of the scale. The Shuijingping case presents two problems for analysis. The first is the issue presented by the Xinzhai data we examined: some of the alternations have lost their phonetic grounding, while remaining structurally identical to the natural alternations from which they are derived. The second problem is equally interesting: following the split, the same set of relationships are retained in each series. However, a conspiracy between time and coincidence has erased the phonetic coherence which once characterized these alternations.

¹⁴Xinzhai also underwent this split and displays the same double chain shift was Shuijingping. However, because of a lacuna in the data, I do not use Xinzhai to illustrate that development.

¹⁵Actually, this set included not just aspirated plosives and affricates, but also “aspirated” fricatives and voiceless sonorants (nasals and liquids).

We are left with an analytic conundrum: if a feature is posited that minimally distinguishes the corresponding members of each series, it would be expected that the two series would behave identically except in processes that make reference to that feature. However, such a feature could have no consistent phonetic correlate: it would be purely diacritic in its function. Furthermore, equally abstract features would be needed to relate HM to LML, \uparrow H to MH and H to ML. On the other hand, we could avoid positing such diacritic features and instead posit rules or constraints that apply separately to the tones in each series. To wit:

- (127) a. \uparrow H \rightarrow H / {M, L, HL}_-
- b. HM \rightarrow \uparrow H / {M, L, HL}_-
- (128) a. MH \rightarrow ML / {M, L, HL}_-
- b. LML \rightarrow MH / {M, L, HL}_-

This approach seems flawed, however, in that it does not capture the parallelism between the two chains. It treats them as completely unrelated processes and does not capture the fact that, in both a historical and a synchronic sense, HM is to \uparrow H as LML is to MH and \uparrow H is to H as MH is ML¹⁶.

A similar phenomenon can be observed both internal to Shuijingping and between Shuijingping and Mashan (and the other Western Hmongic languages discussed above). At its birth, the chain shift could be described entirely in terms of pitch, just as it still can be in Dananshan Hmong and A-Hmao. Through time, however, the tones have wandered in the tone space, leading to phonetic changes without corresponding changes in their phonological behavior. The Mashan chains no longer display scalar properties

¹⁶Just as problematically, it cannot capture the generalization that all of these changes have precisely the same conditioning environment. This appears to be a coincidence, even though it clearly is not. However, discussing the mechanism by which conditioning environments are captured in Structural Optimality is beyond the scope of this chapter.

in the substantive domain. However, in the logical domain they are no less scalar than the Dananshan and A-Hmao chains. Once we assume that our extrinsic scales have no inherent phonetic content, but are, rather, abstract relations over contrasting phonological elements, it is possible to find the face behind the many masks of this protean chain. This type of analysis has already been given above, for the transparent Dananshan chain. Just as the partial scale

$$(129) \quad T = \{M\}_2 > \{H\}_1 > \{\uparrow H\}_0$$

was posited for Dananshan, we posit the partial scale

$$(130) \quad T = \{H, ML\}_2 > \{\uparrow H, MH\}_1 > \{HM, LML\}_0$$

for Shuijingping. Assuming a “diacritic” feature (or trivial scale)¹⁷ that distinguishes the “upper” register from the “lower” register:

$$(131) \quad R = \{LML, MH, ML\}_1 > \{HM, \uparrow H, H\}_0$$

and a SAME constraint that prevents changes along this scale, we can model this pattern of alternations with a grammar that is substantially identical to the one presented for Dananshan. If we call the register scale R , then our grammar would consist of the rankings:

$$(132) \quad \text{HIGHER}[T] \gg \text{SAME}[T] \gg \text{END}[T], \text{DIFF}[T], \text{SAME}[R]$$

¹⁷From a historical standpoint, the Mashan dialects actually have a three-register system. The four historical tones were first bifurcated based upon the voicing of onsets, where syllables having voiceless or preglottalized onsets ended up in one category and syllables with plain voiced onsets ended up in another. This split is shared with most of the other Western Hmongic languages. The second split was based upon aspiration. Syllables with aspirated onsets (including voiceless sonorants) were assigned to a lower register and other syllables were assigned to a higher register. It is to this later split that we make reference here, but all three levels in the register system are synchronically relevant in Shuijingping in that all tone sandhi alternations are “register preserving.” As such, the scale R described below should almost certainly have three levels rather than two.

where SAME[R] is not crucially ranked relative to the other constraints shown here but must dominate HIGHER[R], DIFF[R], and ENDMOST[R].

The resulting grammar will generate both of the relevant chain shifts:

(133)

HM	HIGHER(<i>T</i>)	SAME(<i>T</i>)	END(<i>T</i>)	DIFF(<i>T</i>)	SAME(<i>R</i>)
(a) HM	*!			*	
☞ (b) ↑H		*	*		
(c) MH		*	*		*!
(d) H		*!	**		
(e) ML		*!	**		*

(134)

↑H	HIGHER(<i>T</i>)	SAME(<i>T</i>)	END(<i>T</i>)	DIFF(<i>T</i>)	SAME(<i>R</i>)
(a) ↑H	*!		*	*	
(b) MH	*!		*	*	*
☞ (c) H		*	**		
(d) ML		*	**		*!

(135)

H	HIGHER(<i>T</i>)	SAME(<i>T</i>)	END(<i>T</i>)	DIFF(<i>T</i>)	SAME(<i>R</i>)
(a) ↑H	*	*!	*		
(b) MH	*!	*	*		*
☞ (c) H	*		**	*	
(d) ML	*!	*	**		*

(136)

LML	HIGHER(<i>T</i>)	SAME(<i>T</i>)	END(<i>T</i>)	DIFF(<i>T</i>)	SAME(<i>R</i>)
(a) LML	*!			*	
☞ (b) MH		*	*		
(c) ↑H		*	*		*!
(d) ML		*!	**		
(e) H		*!	**		*

		MH	HIGHER(<i>T</i>)	SAME(<i>T</i>)	END(<i>T</i>)	DIFF(<i>T</i>)	SAME(<i>R</i>)
(137)	(a) MH		*!		*	*	
	(b) ↑H		*!		*	*	*
	↗ (c) ML			*	**		
	(d) H			*	**		*!
		ML	HIGHER(<i>T</i>)	SAME(<i>T</i>)	END(<i>T</i>)	DIFF(<i>T</i>)	SAME(<i>R</i>)
(138)	(a) MH		*	*!	*		
	(b) ↑H		*!	*	*		*
	↗ (c) ML		*		**	*	
	(d) H		*!		**	*	*

Not only that, it can generate all of the tonal chain shifts we have observed thus far (provided only that the tones in the scales are changed for their cognates).

3.2.3.1 Naturalness, productivity, and the proper scope of phonology

There are a number of possible objections that could be raised to an analysis of this sort. These fall into phenomenological and methodological categories. On the phenomenological side, it is perfectly reasonable to question whether speakers possess any productive knowledge of the highly abstract alternations we have been discussing here. Indeed, some scholars have suggested that this type of tone sandhi consists almost entirely of fossilized historical patterns stored in the lexicon and not productively commanded by speakers (Ballard 1988). On the methodological side, it must be decided whether phenomena of this kind should be treated as phonology at all.

One might attribute these phenomena to some mechanism outside phonology. In this vein, Tsay and Myers (1996) dismissed Taiwanese tone sandhi as “allomorph selection.” Under such a model, the selection of normal versus sandhi allomorphs would be directly equivalent to the mechanism by which *a* versus *an* are chosen in English.

For each form participating in the sandhi alternation there would be two allomorphs: one for the normal form and one for the sandhi form. The allomorph selection rule would then choose an allomorph based on the phonological environment of the morpheme. This mode of explanation can easily account for much of the data seen here. It is problematic, though, in that it can account for almost all of lexical phonology and a large part of postlexical phonology. In fact, it is possible to view any model of phonology that attempts to account for alternations as an allomorph selection mechanism. If this pattern is to be dismissed as “allomorph” selection with little or no relevance to phonology, a principled explanation should be offered to differentiate this direct allomorph selection—presumably a function of the morphology—from “real phonology”.

One possible criterion is naturalness. On this view, phonetically natural phonology is phonology proper; phonetically aberrant “phonology” belongs in the domain of morphology. This argument presents two major problems, one general and one specific.

The general problem is that confining phonology in this way invariably leads to a kind of methodological circularity. As practitioners of a typological discipline, phonologists seek to discover the range of possible phenomena in the sound patterns of human languages. This endeavor becomes an exercise in tautology if it is limited by subjective criteria such as naturalness. To circumscribe phonology with an arbitrary naturalness requirement would be to practice typology by fiat. Put more bluntly, it is pointless to seek the range of possible phonological grammars when the boundaries of that range are definitionally predetermined. In fact, the allomorph selection hypothesis seriously erodes the restrictiveness of morphophonological theory since it predicts that any allomorphy should be possible.

This global problem is closely related to a more specific problem raised by the data discussed here. A-Hmao, and especially Dananshan, look relatively “natural” while the Mashan cases look less so. However, even in these cases a bit of the original low-

ering pattern is preserved. To exclude the Mashan data on account of unnaturalness would require a criterion that defined the cut-off point between “natural” and “unnatural” phonology, but wherever it was set, such a boundary would disguise the fact that “naturalness”—in so far as it can be given a principled definition of any kind—is a gradient, rather than a categorical, concept. On these grounds, any arbitrary naturalness criterion should not be invoked to exile these data from the realm of phonology.

Another (more principled) objection that can be raised is that of productivity. It must be granted that the tone sandhi processes that have been discussed here vary in productivity. In Dananshan, for example, the tone sandhi process is not particularly productive¹⁸ (Wang 1985). In A-Hmao, the process seems to be much more productive, as suggested by Downer (1967) and the text provided by Wang (1986). In the Mashan dialects the tone sandhi appears to be quite productive, although Xian (1990) is inexplicit on this point. The even more surprising tone sandhi alternations found in Southern Min languages (for which, see §4.1) are completely productive, excluding certain classes of loanwords. Interestingly enough, there does not seem to be a clear correlation between the phonetic transparency of the tone sandhi processes and their productivity, at least in the languages represented here. Even granting, for the sake of argument, that none of these processes is completely productive, it is not clear that these processes should be excluded from phonology proper. Indeed, a great deal of lexical phonology is not productive or is productive only in a specific subset of the lexicon. If the tone sandhi processes discussed here are to be denied a place in the domain of phonology, then much of what has been called phonology in the past must join them in their exile.

What criteria, then, can be offered to distinguish phonologically conditioned allomorph selection processes from phonology proper? Responsibility for an allomorph

¹⁸Wang (1985) notes specifically that these processes do not occur in a stratum of recent loans from Chinese, though they do occur in older loans. See also Niederer (1998).

that is unambiguously suppletive must clearly be delegated to the morphology. Furthermore, processes that apply to an arbitrary subset of the lexicon (e.g., a single lexical item as in the case of English *a/an*). All of the chain shifts we have examined can be derived by the application of very simple (if “unnatural”) rules and apply to a large, non-arbitrary subset of the lexicon of the languages in question. In other words, it is right to attempt to example them with the same set of tools used to explain the rest of phonology, and if necessary, to enrich the set of explanatory devices employed in phonological analysis such that it able to account for them.

The Western Hmongic tonal shifts are historically and structurally coherent but (in the case of the Mashan dialects) phonetically arbitrary. We have argued that they are nevertheless phonological phenomena, and have argued for a formal mechanism able to account for these shifts in such a way as to capture their underlying similarity while allowing for their phonetic divergence. These formal mechanisms are abstract phonological scales and constraints referring to these scales. Given formal mechanisms of this power, it is now possible to account for other problematic chain shifts.

3.3 Fricative Place Chain Shift in Acquisition

As mentioned above, there is fricative place chain shift in the speech of many children acquiring English. This shift is described in detail by Dinnsen and Barlow (1998), who provide several different analyses of the phenomenon, approaching it from different perspectives. This chain shift has the effect of moving voiceless fricatives one step “forward” in the articulatory space so that /s/ is realized as [θ] but both /θ/ and /f/ are realized as [f].

After examining several different analyses (some of which treat this process as a chain shift and some of which do not) they settle on accounts based on “shadow spec-

ification,” a type of underspecification intended to allow the same representations to function as a basis for production and perception in language development. In their Optimality Theory analysis using shadow specification, *thin* and *fin* have the same underlying representation /Fɪn/, while *sin* is represented underlyingly as /sɪn/. This implies that learners displaying this pattern do not distinguish perceptually between [θ] and [f], but do distinguish between these sounds and [s]. The grammar effectively forces /F/ to surface as [f] and /s/ to surface as [θ], meaning that—under this analysis—the phenomenon is not a chain shift at all. They fail to specify, however, what happens to representations like underlying /θ/, which presumably would surface as [θ] or underlying [s] which would apparently surface as [s]¹⁹.

In contrasting this analysis with the analysis attributed to them above, a chain-shifting analysis in which the underlying representations are fully adult-like, the three places of articulation are treated as a scale, and too great movement along this scale is penalized by DISTFAITH. They reject this analysis on several accounts. First, they express doubt that the scale they need to posit (/f/ = 1, /θ/ = 2, /s/ = 3) can have any real basis in phonology. Secondly, they point to the fact that the subjects, after speech therapy, show a pattern of lexical diffusion, where some lexical items come to be realized with adult-like values for these sounds and other words retain the chain-shifted realizations. They discount, out of hand, the possibility that this lexical diffusion is due to the coexistence of competing, coexistent grammars (what have been called “cophonologies”). However, a mechanism of this kind is independently motivated for the lexical stratification found in adult speech (see, e.g. Itô and Mester (1995, 1999) for the case of Japanese), so Dinnsen and Barlow’s rejection of that explanation is at best premature. They also note the questionable appeal to the otherwise unmotivated constraint

¹⁹This is consistent, in some sense, with the data from acquisition in that learners who display the chain shift pattern, after therapy, may display lexical variation such that adult /s/ sometimes corresponds to [θ] and sometimes to [s]. Likewise with adult /θ/, *mutatis mutandis*.

DISTFAITH. This problem (and its resolution) has been discussed above.

There are some respects in which the first analysis given by Dinnsen and Barlow (1998) is superior to the first. One is that of economy—the first explanation requires no special assumptions about underlying representations, which may be identical to adult representation. This mirrors the general observation that children’s perceptual development leads their articulatory development; that is, they can perceive distinctions before they can produce them. The only additional representation device which is needed is the scale. More significantly, the first account treats the chain-shifting process as a unified phenomenon, where both steps are driven by markedness. The second analysis makes the two steps of the chain shift completely disjoint. The first step, they attribute to greater markedness of /s/ relative to /θ/²⁰ The second step, they attribute to the absence of an underlying contrast. At the expense of eliminating the appearance of opacity, they have been forced to give disparate analyses to similar phenomena.

Based upon what we have already seen about the relationship between phonological scales and the substance of the items they classify, the existence of a scale with the form

$$(139) \quad \{f\}_2 > \{\theta\}_1 > \{s\}_0$$

is no longer surprising (and comes at no additional theoretical cost). Indeed, this scale looks relatively grounded compared to the tone scales we discussed above. Structural Optimality provides a way of capturing the insight that this chain shift is a unified phenomenon and avoids making the prediction that the affected children perceive no contrast between /f/ and /θ/ (but do perceive a contrast between /θ/ and /s/) without employing DISTFAITH or any other chain shift-specific theoretical mechanism. Given the ranking schema we have already developed, and the scale from (139), the chain shift above follows naturally, as shown below (140-142):

²⁰This, in itself, seems to be a rather strange state of affairs since, cross-linguistically, the opposite pattern generally holds.

		/sm/	HIGHER(<i>P</i>)	SAME(<i>P</i>)	END(<i>P</i>)	DIFF(<i>P</i>)
(140)	(a) sm		*!			*
	(b) θm			*	*	
	(c) fm			*	**!	
		/θm/	HIGHER(<i>P</i>)	SAME(<i>P</i>)	END(<i>P</i>)	DIFF(<i>P</i>)
(141)	(a) sm		*!	*		
	(b) θm		*!		*	*
	(c) fm			*	**	
		/fm/	HIGHER(<i>P</i>)	SAME(<i>P</i>)	END(<i>P</i>)	DIFF(<i>P</i>)
(142)	(a) sm		*	*!		
	(b) θm		*!	*	*	
	(c) fm		*		**	*

The theory developed to account for the tonal chain shifts in Western Hmongic predicts as a side-effect the possible existence of consonant place chain shifts. This case shows that such effects do in fact exist and call for the same type of analysis.

3.4 Conclusion

We have shown how the theoretical mechanisms of Structural Optimality account for the chain shifting phenomena as unified processes. This is not a novel accomplishment, since—as we have seen—earlier investigators have been able to encode much the same insight in OT grammars, albeit by different means. However, this proposal has accomplished several things that earlier proposals have not. Significantly, it has provided an analysis of chain shifts in terms of a mere four constraint types, all of which are independently motivated. As will be seen, these same types of constraints and scales are needed to account for circle shifts, bounce-back effects, ordering effects in coor-

dinate compounds, long distance dissimilation effects, and other phenomena. Just as significantly, we have been able to capture—in a way that none of the previous theories would be able to do—the logical relation that persists when the winds of diachrony sweep away the phonetic naturalness that characterizes young chain shifts without destroying their underlying structure.

Chapter 4

Circular Chain Shifts

I see how plenty surfeits oft,
And hasty climbers soon do fall;
I see that those which are aloft
Mishap doth threaten most of all.

“My Mind to Me a Kingdom Is”
Sir Edward Dyer

Questions about circular chain shifts (hereafter *circle shifts*)—whether they exist and, if so, how they are to be analyzed—have become a prominent issue in phonological theory (McCarthy 2002; Łubowicz 2003; Hsieh 2004; Zhang 2006; Barrie 2006). The increased interest in these processes is intimately tied to certain formal properties of Optimality Theory. As classically defined, Optimality Theory predicts that phonological circle shifts should not exist, as was demonstrated irrefutably in a formal proof constructed by Elliot Moreton, a finding which will be referred to hereafter as *Moreton’s Theorem* (Moreton 2004b).

This chapter will examine four different cases that have been described as circular chain-shifts, namely the tone circles in Southern Min (Chinese), Western and Eastern A-Hmao (Hmongic), and Jingpho. I will demonstrate that the theoretical equipment that is necessary to model these circle shifts (and “bounce-back effects”) is exactly the same as that required to model the conventional chain-shifts examined in Chapter 3, the coordinate compound ordering effect to be examined in Chapter 5, and the graded-dissimilation effects in reduplication to be analyzed in Chapter 6. This chapter will be a first foray into to the domain of anti-faithfulness, wherein empirical arguments for

both anti-identity and directional anti-identity will be presented. These same themes will become crucially important in Chapters 5–6, where it will be shown that the input-output anti-identity constraints needed for analyzing circular chain shifts and bounce-back effects are simply a special case of correspondence constraints that must also exist in string internal forms. This unified case for anti-identity, and its predictions for phonology, will be shown to compare favorably with analyses of circle shifts that rely upon contrast preservation, or which try to deny their existence altogether.

While there has been a long debate (Wang 1967; Hsieh 1970; Cheng 1968, 1973; Moreton 2004b; Barrie 2006; Zhang 2006) about the nature and proper analysis of circle shifts, this debate has continually been mired in the same evidential swamp since its beginning. One of the reasons so little progress has been made into resolving this issue—perhaps the principle reason that this is so—is that the debate has centered wholly around the analysis and reanalysis of the same circle shift, namely the (justifiably) famous tone sandhi ring of Xiamen and Mainstream Taiwanese (two closely related dialects from the Southern Min branch of Chinese). As long as this is so, it is unlikely that a general understanding of circle shifts and their properties will ever emerge. This chapter is an attempt to address that missing piece of the scholarly puzzle. I will show that Structural Optimality can account neatly for a variety of types of circle shifts.

Our discussion will begin, in §4.1, with a discussion of the Min Tone Sandhi circles, both the famous circle from Xiamen and Mainstream Taiwanese in §4.1.1 and striking variety of other Southern Min tone circles in §4.1.2–4.1.4. It will be seen that these latter shifts are not only amenable to Structural Optimality and intractable for contrast-based theories of tone circles, but also fill out gaps in the typology of chain shifts predicted by the basic architecture of our theory. We will also discuss the historical origins of these patterns and ask what this means for synchronic analyses of this

type of tone sandhi. It is significant, of course, that circle shifts are not confined to Southern Min. We will then explore in-depth analyses of circle shifts (bounce-back effects) in A-Hmao (§4.2) and Jingpho (§4.3) these cases help to answer the concern, expressed by a number of phonologists, that circle shifts do not occur outside of the Min dialect group in Sinitic (Schuh 1978; Moreton 2004b). They also present striking confirmations for a number of mechanisms within Structural Optimality, including both general and directional anti-identity.

4.1 Southern Min Tone Sandhi Circles

The Southern Min dialects of Chinese are known for their complex systems of tone sandhi. In these languages, every tone typically has a sister tone that surfaces in sandhi environment. The sandhi environment, for most of the dialects in this group, at least, is very general: the base tone occurs in isolation and phrase finally; the sandhi alternate appears when the syllable bearing the tone is non-final. These alternate tones may either be identical to one of the base tones (structure-preserving, so to speak) or may be distinct from any of the isolation tones (structure-building, contrast-preserving). We will refer to the general class of tones that appear in the sandhi context as *sandhi tones* and the tones that appear only in that context as *pure sandhi tones*. This distinction will become important in our subsequent discussion, in both its synchronic and diachronic dimensions.

As we have already discussed, the concentration of the attention of phonologists upon the Taiwanese tone sandhi circle belies the great richness of circle shifts that appear in other Southern Min dialects. Some of these share a common history with the Taiwanese circle; others are quite independent. Both types of circles are useful in evaluating theoretical claims that have been made about tone circles and in arriving at a more adequate theory of this phenomenon. These other circles have particular relevance

to the claims recently made by Hsieh (2004) and Barrie (2006) that tone circles are driven by the need to avoid marked structures while also avoiding neutralizations. A few tone circles, however, cannot be explained this way due to the fact that, for these patterns, the input inventory is identical to the output inventory. Before we progress to these cases, however, it is my obligation to introduce the better-known tone sandhi circle of Mainstream Taiwanese and some dialects from around Xiamen (Amoy).

4.1.1 Xiamen and Mainstream Taiwanese

Mainstream Taiwanese has seven “tones.” Specifically, it has a two-way tonal contrast in *checked syllables* (syllables with an obstruent coda) and a five-way tonal contrast in *smooth syllables* (syllables without an obstruent coda), as shown in (1):

- | | | | |
|-----|----------------|-----|--------------|
| (1) | <i>smooth</i> | 24 | high rising |
| | | 22 | low |
| | | 21 | low falling |
| | | 53 | high falling |
| | | 44 | high |
| | <i>checked</i> | 32q | low checked |
| | | 54q | high checked |

The general rule for tone sandhi alternations in Xiamen is as follows:

- (2) $T \rightarrow T' / \text{“non-finally”}$

The specific alternations are given in (3):

- (3) a. Rising and **high** become low.
 b. Low becomes low falling.
 c. Low falling becomes high falling.
 d. High falling becomes **high**.

Thus, there is a complete “chain” leading from high back to high. Data illustrating these alternations (taken from Chen (2000)) are given in (4):

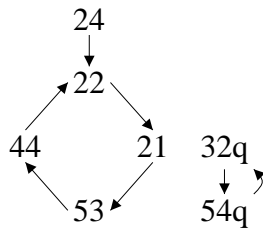
- | | | | | | |
|-----|----|-------|------------|------------|---------------------------------|
| (4) | a. | p'ang | 'fragrant' | p'ang tsui | 'perfume' ('fragrant'+ 'water') |
| | | 44 | | 22.53 | |
| | b. | we | 'shoes' | we tua | 'shoelaces' |
| | | 24 | | 22.21 | |
| | c. | pī | 'sick' | pīlang | 'patient' ('sick'+ 'person') |
| | | 22 | | 22.24 | |
| | d. | ts'u | 'house' | ts'u ting | 'roof' ('house'+ 'top') |
| | | 21 | | 53.53 | |
| | e. | hai | 'ocean' | hai kī | 'ocean front' |
| | | 53 | | 44.24 | |

In some dialects, there is a second circle, a simple toggle, between the checked-syllable tones (Wang 1967; Cheng 1983; Ballard 1988; Moreton 2004b):

- (5) a. Low falling becomes high falling.
 b. High falling becomes low falling.

As in Chapters 2 and 3 above, I will employ the convention of using directed graphs to represent interrelated phonological alternations (i.e. neutralizations, chain shifts, and circle shifts). The graph for Taiwanese tone sandhi is given in (6). The tones are represented, following the convention in discussions of this phenomenon, using Chao tone numbers. The post-posed letter *-q* is used to designate tones occurring in checked syllables.

(6) **The Xiamen/Mainstream Taiwanese tone sandhi circle**



This graph represents one of the subvarieties of Taiwanese in which there is a second circle between the two tones that occur in checked syllables (the *rusheng* tones). This is a type of Taiwanese that we will consider in our initial analysis, making allowance for some of the earlier analyses that were based on tone systems of other subvarieties in which both *rusheng* tones have pure sandhi tones as the sandhi-context alternates.

4.1.1.1 A digression on “psychological reality”

Outside of Southern Min dialects, there is no other language known to have precisely this type of circular chain shift. In fact, many investigators still doubt the existence of circular chain shifts as phonological phenomena (McCarthy 2002; Łubowicz 2003; Moreton 2004b), given the fact that relatively few of these cases have been reported. The great rarity of phonological patterns of this kind, combined with the great difficulties that have presented themselves when phonologists have attempted to model these phenomena with mainstream theories have motivated attempts to remove the infamous tone circle from the range of phenomena for which phonological theory is responsible to account.

One of the earliest and most convincing arguments against treating the tone circle, and other related tone sandhi patterns, as tonal phenomena that were distinct from the tonal phonology seen in other languages, where tonal alternations could be elegantly modelled as featural changes, was to label them “paradigmatic replacement” Schuh (1978). Without denying the reality of the phenomenon, Schuh asserted that it was something unique—a phenomenon that was not particularly useful in understanding the

features of tones and the of structure tonal processes as seen elsewhere. A not-unrelated position has been advanced by a number of other investigators, notably Ballard (1988) for Southern Chinese dialects and Downer (1967) for Hmongic, is that tone sandhi alternations of this kind are really the relics of sound changes that have occurred in the history of these tone systems. Under this point of view, these tonal alternations are neither well understood in terms of synchronic featural analyses nor particularly useful in understanding tonal alternations elsewhere.

A more common objection to treating the tone circle as part of phonology proper—really a rephrasing of Ballard’s position—is the assertion that the alternations described in this case are not “psychologically real” and should therefore have no bearing on theories of synchronic phonology. This argument rests upon two claims which I argue to be overbold: first, that the psychological unreality of the Southern Min tone sandhi alternations has been satisfactorily demonstrated and second, that psychological reality—whatever that may mean—is the proper criterion for deciding what phonological theory need and need not explain.

Most attempts to test the psychological reality of Xiamen/Mainstream Taiwanese tone sandhi have relied upon wug-tests (Berko 1958). In these tests, speakers were presented with nonce forms and asked to either compose phrases out of them or to recover individual morphemes from a disyllabic phrase. In these tests, the participants uniformly showed little accuracy in either producing or reversing tone sandhi in nonce forms, sometimes doing no better than chance. These results have been held to demonstrate that the tone sandhi patterns in Taiwanese simply represent fossilized allomorphy with no status in the synchronic grammar. This assertion, however, is probably premature. At the most basic level, it is unclear that the “wug tests” that have been performed to test Taiwanese tone sandhi do not seem to be comparable to wug tests of the type devised by Berko (1958). In these tests, children were asked to apply a particular mor-

phonological operation to a base. Because the semantics of the base were made clear, and because the participants were prompted to apply an operation to the base that could not be applied to utterances of other types, participants had a great deal of information about what the nonce form they were dealing with was supposed to be (a singular noun in the participants native language). It is not clear that the same level of information was available to the participants in the Taiwanese tone tests, to the extent that the way that the participants were treating the stimuli as tokens of a rather different type than that intended by the experimenters, perhaps not even as intended Taiwanese words (regardless of whether they were given direct instruction that this was the case). It is therefore not clear that the results of these tests are at all comparable to Berko's wug tests, nor is it clear that passing or not passing such a test is a reliable criterion for establishing "psychological reality."

This is due, in no small part, to the fact that "psychological reality" is not a well-defined concept. There is no experimental procedure that allows the experimenter to plumb the depths of the human mind—we are limited to observations of behavior. By the same token, while it is clear that the ability to apply an old generalization to new forms shows that the generalization is part of that individual's linguistic competence, there is no solid basis for saying that the refusal of speakers to extend such a generalization to a particular set of forms presented to them by an experimenter indicates that those speakers have no knowledge of that generalization or an ability to employ it in language use. A more humble approach to the subject is to say that we do not know what knowledge and cognition is driving linguistic behavior, but only what the patterns in that behavior are.

The argument, then, is that "psychological reality" should play no role in determining what phonological theory, as such, is responsible to explain¹. In practice, in

¹This is not to say that the psychological aspects of the sound patterns of languages are uninteresting or unworthy of examination. Rather, it is an argument that phonological theory driven by the classical

fact “psychological reality” is usually invoked as a means of dismissing a pattern or phenomenon for which one’s theory of choice is unable to account. One of the best examples of this is the subject at hand, namely the Min tone sandhi circles. Because of the selective way in which this criterion is typically applied, it has a deleterious effect upon the development of phonology. Investigators, presented with two generalizations, A and B, where A is compatible with existing theory and B is not, are likely to investigate the psychological reality of B (e.g. Min tone sandhi) rather than A. The result of this practice is simply the reinforcement of existing ideas in the face of potentially disconfirmatory facts. Aside from the fact that they may mirror, in some respects, the linguistic behaviors of thinking humans, there is little reason to believe that any existing formal theory of phonology bears any special relationship to cognition. If these endeavors are rightly-guided, the psychological reality of the patterns of linguistic behavior, either that predicted by a formal grammar or produced by a human, is not of crucial interest in formulating a formal theory of phonology.

4.1.1.2 Allomorph-selection hypotheses

Of course, if one argues that Southern Min tone sandhi does not represent a psychologically real set of phonological processes, one nevertheless must account in some way for the existence of these productive alternations. The most widely repeated of these explanations is the “allomorph selection hypothesis” which seeks to place the locus of Xiamen/Mainstream Taiwanese tone sandhi in the morphology rather than the phonology. Early echoes of this idea are to be found in Hsieh (1970, 1976) and Ballard (1988). However, the best known expression of this hypothesis is Tsay and Myers (1996). This type of hypothesis has the advantage of freeing the phonology from the burden of generating these highly “unnatural” and unusual patterns. It appears to be consistent, too,

linguistic trope of expressing language patterns at the highest possible degree of generality is a discipline of value in its own right, regardless of the psycholinguistic concomitants of these patterns.

with the relatively poor performance of Taiwanese speakers in the wug tests criticized above.

However, as pointed out by recent investigators like Hsieh (2004) and Barrie (2006), this approach to the problem has difficulties of its own. First, it fails to account for why certain classes of loanwords participate in the process. If the pattern was simply fossilized allomorphy resulting from sound changes that occurred in the histories of these languages, there is no good reason to expect that it should apply to any loanwords at all.

There is a more fundamental philosophical problem with these arguments, though. In saying that these alternations are simply allomorphy, with no relationship to the phonology, we miss the crucial generalization that, in sandhi context, syllables bearing one tone in the input share the same tone in the output, which always differs from their shared tone in the input. At least since the neo-grammarians, the central trope of linguistics, and especially phonology, has been to identify all that is systematic about language, and find the most general way of capturing this systematicity, leaving behind only the irreducible core of a language, a language family, or a synchronic grammar. In adopting a loose psychologism as the basis for deciding what patterns a theory of phonology should account for and what patterns a theory of phonology should ignore, we undermine and abandon the whole traditional enterprise of western linguistics. This is not to say that an understanding of the psychology of language is not important—it is useful and interesting in its own right. However, considerations of this type should not free phonological theory from accounting for regularities and subregularities in the grammars and lexicons of language.

Of course, it is difficult not to believe that the phenomenon involved in Southern Min tone sandhi involves the lexicon in some way, or even that it is an type of allomorph selection. Indeed, by definition, the same is true of all morphophonology. Within a lan-

guage, there exists a finite set of “allomorphs” for each “morpheme” (to the extent that the morpheme is a meaning construct; see Anderson (1992)). Phonological grammars, especially during the generative period, have been intended to choose which of these allomorphs occur in certain environments by the application of principles—rules or constraints—that are general to all forms sharing particular phonological properties (or, to be more realistic, all forms within a particular morphological or lexical subclass that share particular phonological properties. Southern Min tone sandhi is completely regular within its lexical domain (that is, within all words but a specific class of recent loans from English); thus, a theory of morphophonology that cannot account for this set of alternations generally, rather than simply as a list of pairs may be psycholinguistically realistic, but is descriptively inadequate.

4.1.1.3 Rule based analyses

The earliest generative analyses of Xiamen/Mainstream Taiwanese tone sandhi were formulated in terms of SPE-style rewrite rules. On the face of it, there are three ways that one could generate a circular chain shift like that in Xiamen and Mainstream Taiwanese using rules of this type:

- (7) a. **Minus-alpha notation.** Feature-switching could allow the generation of these patterns.
- b. **Build only to destroy.** A non-structure-preserving rule could change the representation of one item to a new category prior to the application of rules implementing a chain shift; following the application of the chain-shift rules, this new category must be eliminated in favor of the category to which which the next item on the chain formerly belonged. To wit:

- i. $A \rightarrow A'$ ii. $B \rightarrow A$ iii. $C \rightarrow B$ iv. $A' \rightarrow C$

- c. **Multiple representations for one surface form.** The “beginning” of the circle and the “end” could be given different underlying representations, even though their surface forms are identical.

The first—and most famous—of these was the analysis proposed by Wang (1967). This extremely clever analysis generates the whole circle using a single, rather simple, alpha-switching rule:

$$(8) \begin{bmatrix} \alpha \text{ HIGH} \\ \beta \text{ FALLING} \end{bmatrix} \rightarrow \begin{bmatrix} \beta \text{ HIGH} \\ -\alpha \text{ FALLING} \end{bmatrix}$$

This allowed Wang to generate the following sequence:

$$(9) \begin{bmatrix} +\text{HIGH} \\ -\text{FALLING} \end{bmatrix}_{44} \rightarrow \begin{bmatrix} -\text{HIGH} \\ -\text{FALLING} \end{bmatrix}_{22} \rightarrow \begin{bmatrix} -\text{HIGH} \\ +\text{FALLING} \end{bmatrix}_{21} \rightarrow \begin{bmatrix} +\text{HIGH} \\ +\text{FALLING} \end{bmatrix}_{53} \rightarrow \begin{bmatrix} +\text{HIGH} \\ -\text{FALLING} \end{bmatrix}_{44}$$

However, this analysis is problematic in a number of respects. First of all, as pointed out by Chen (2000), there are topologically identical tone circles—which are clearly related historically to the Mainstream Taiwanese circle—for which this analysis cannot work. The reason for this was a theme of Chapter 3 which will also be fundamental to many analyses in Chapter 5: tones tend to wander in phonetic space, but this wandering does not necessarily change their structural relation to one another. The fact that the rule proposed by Wang can only account for the one case, in the face of other cases that have identical properties aside from the phonetic ones, is something of an embarrassment, and draws attention to the very contrived nature of this analysis.

The final option was taken by Yip (1980) and several other investigators. In Yip’s analysis, [33] has two different representations. This has the effect of reducing the circle to a simple chain shift. Moreton (2004b) and Barrie (2006) both contend that this analysis seems ad hoc. However, it is justified on theory-internal grounds that tonal inventories should contain two mid tones (IH and hL). Furthermore, other analyses in which two underlying representations must have two surface realizations have been

advanced in work such as Hyman (1988). Furthermore, a similar concept is indispensable to the analyses of White Hmong (Ratliff 1992b) and Eastern A-Hmao (as seen in §3.2.2). However, the arguments by Moreton and Barrie that Yip's analysis is somewhat arbitrary are not without merit—Yip requires, for example, that both IH and hL be realized as 22. While this approach is consistent with the kind of abstraction away from phonetics that is advocated in the current study, it does not appear to be compatible with the general assumptions of the type of autosegmental approach adopted by Yip.

The greatest problem with all of these analyses, however, is the fact that they are not truly explanatory. While they allow us to model these patterns within a certain framework, they provide no enlightenment as to why such patterns should exist or what the fundamental nature of these patterns are.

4.1.1.4 Contrast preservation analyses

Efforts to address this lack of explanatory adequacy have come in the form of analyses that depend upon the notion of contrast preservation. The basic idea of these analyses is that grammars maintain a balance between reducing the markedness of an inventory and preserving underlying contrasts in the output. Such analyses are difficult or impossible to implement in a meaningful way using rule-based formalisms. However, Optimality Theory (especially certain variants of OT; see Łubowicz (2003); Flemming (2004)) has proved particularly well-suited for expressing analyses of this kind. Such analyses have the great advantage that they can motivate such alternations through two general principles.

A very interesting example of such an analysis is presented by Hsieh (2004), who models chain shifts in terms of contrast preservation. To model circular chain shifts, he uses both contrast preservation and output-output antifaithfulness. As originally formulated, the theory of contrast preservation presented by Łubowicz (2003) predicted

that circular chain shifts should not exist. By adding transderivational antifaithfulness, motivated by a need for paradigmatic contrastiveness, Hsieh was able to model the Mainstream Taiwanese tone sandhi circle. In order for this model to work, it appears that Hsieh needed both contrast preservation and antifaithfulness in his model.

Barrie (2006), however, presents a model of the Mainstream Taiwanese tone sandhi circle which relies only upon contrast preservation—a very slightly modified version of Łubowicz’s (2003) contrast preservation formalism. In this formalism, EVAL (the function corresponding to the ranked-constraint grammar of OT) is divided into two stages: one in which PC CONSTRAINTS (preserve contrast constraints) and markedness apply, which she (Lubowicz) calls H-eval₁, and one where generalized faithfulness applies, her H-eval₂. Note the following architectural illustration taken from Łubowicz (2003):

(10) **Structure of PC grammar**

- a. Gen(In_k)
- b. H-eval₂(H-eval₁(Scen_i ≤ $i < \infty$))

Where:

H-eval₁ = PC and Markedness

H-eval₂ = Generalized Faithfulness

These grammars do not evaluate individual input-candidate pairings; instead, they evaluate “scenarios” which are relations corresponding to a whole system of input-output mappings. Owing in large part to this novel reworking of the OT framework, markedness and faithfulness in Contrast Preservation Theory are somewhat different from their counterparts in normal OT. Most crucial for Barrie’s analysis (and therefore, for our purposes) is the notion of tokenized markedness. Łubowicz (2003) defines this as follows:

(11) **Tokenized Markedness**

Assign a violation mark for every instance of output, out_x , where the number of outputs equals the number of inputs that map onto out_x . “Assign a violation mark for every token of a marked output in a scenario, where the number of tokens equals the number of inputs that map onto this output.”

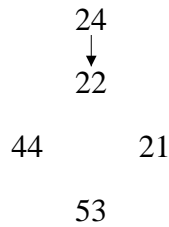
Importantly, Tokenized Markedness constraints make reference to both input and output (rather than just output, as in Classical OT) and penalize marked outputs based upon the number of kinds of inputs that are mapped to them. Thus, if there is a markedness constraint against rising tones (call it *RISE) and only one input /LH/ is mapped to LH, then the resulting scenario would incur only one violation of *RISE. However, if both /H/ and /LH/ were mapped to LH, then the scenario would incur two violations of *RISE. The notion of generalized faithfulness is perhaps less exotic: faithfulness constraints of this kind are different from classical faithfulness constraints only in that they are generalized so that there is only one type of faithfulness rather than several.

Lubowicz (2003:145) makes the claim that her Contrast Preservation Theory (PC Theory) cannot generate circular chain shifts:

To sum up, in PC theory there is no movement unless it improves on PC or markedness. This shows that in PC theory, circular shifts are ruled out in favor of non- circular mappings

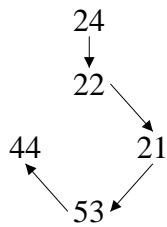
However, Barrie (2006) appears to show that this is not the case, in the process of giving an analysis of the Xiamen/Mainstream Taiwanese tone sandhi circle. In his analysis, there is a high-ranked constraint against rising tones, *RISE. This forces “movement” away from the /24/ rising tone:

(12) **Mapping away from rising tone /24/**



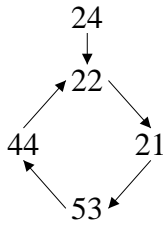
The rising tone /24/ is mapped to [22] because there are also markedness constraints against high tones (that is, [44]) and contour tones ([21] and [53]) so [22] is the ideal output. However, output oriented constraints on register penalize scenarios in which the contrast between /24/ and /22/ is neutralized, so /22/ is mapped to [21]. For the same reason, /21/ is mapped to [53] and /53/ is mapped to [44]:

(13) **Chain shift mapping**



This mapping, however, incurs two violations of the tokenized markedness constraint against high tones (*HIGH) because both /53/ and /44/ are mapped to the high tone [44]. The winning candidate is superior to this candidate in that it incurs no additional violations (both of them involve neutralizations of pitch contour) but the circular mapping avoids the situation where two inputs are mapped to [44] (thus violating *HIGH twice). Thus, the scenario that is selected appears to be the one attested in Mainstream Taiwanese and Xiamen:

(14) **Circle shift mapping**



This analysis has much to recommend it: it is clean, elegant, relies on independently motivated principles, and is predictive.

Unfortunately, however, one of the principle predictions made by this model, and most other theoretical constructs like it, appears to be incorrect. As Barrie (2006) notes, the model he proposes predicts that circle shifts will only occur when there is neutralization somewhere in the system. Chain shifts in Łubowicz’s (2003) framework must always be push chains and must be motivated by a high-ranked markedness constraint that motivates some “movement” within the system. The chain-shifting scenario is selected because it allows markedness reduction while preserving the maximum number of contrasts. If a circle shift involved no neutralization, there would be no markedness reduction in the system and there would, therefore, be no reason for the circle shift to occur in the first place.

Problematically for these theories, however, circle shifts with no neutralization do, in fact, exist. Even within certain dialects of Taiwanese that are closely allied to Mainstream Taiwanese, there are simple exchanges between the tones in checked syllables (*rusheng* tones) with no neutralization involved. However, circle shifts of this kind are very widely attested within Southern Min. Several cases of circles of this kind are documented in §4.1.2. The most spectacular example of this type of shift, however, is Yilan dialect (a Southern Min dialect from Taiwan) which is discussed in §4.1.4. Dongshan dialect, which is discussed in §4.1.5, presents another interesting challenge to contrast preservation theories in that it displays neutralization as *two* different places in the chain, a state of affairs that Łubowicz (2003) specifically predicts to be impossi-

ble. After looking at these other cases which provide very compelling evidence against contrast preservation theories, we will perform an analysis of Xiamen/Mainstream Taiwanese tone sandhi within Structural Optimality.

4.1.2 Checked-syllable (*rusheng*) circles

As Ballard (1988) notes, the circle shift that is most widely distributed throughout the Southern Min family is an exchange between the two tones that occur in checked syllables. The distribution of this type of pattern is indeed extremely wide and must be held, for a variety of reasons, to have developed at least twice (and probably a few times) independently. The reasons that this has occurred will become more clear when we discuss the historical facts that led to the development of circle shifts; however, a brief explanation here is in order. Circle shifts are the descendants of ordinary chain shifts. A circle shift results when two points in the chain merge with one another, producing a loop. For example, suppose that there is a chain shift where $A \rightarrow B$ but $B \rightarrow B'$. If B' then merges with A , it will then be the case that $A/B' \rightarrow B$ but $B \rightarrow A/B'$. This sort of occurrence was especially common in checked syllables since the inventory of tones is greatly reduced in this environment, facilitating mergers between base tones and sandhi tones in a way that they are not facilitated in smooth syllables.

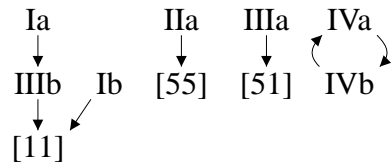
We will look at four rather different cases of *rusheng* circle shifts: those in Dong-an §4.1.2.1, Taiwanese dialects §4.1.2.2, Chaoyang §4.1.2.3, and Chao-an §4.1.2.4.

4.1.2.1 Dong-an

In the Southern Min dialect spoken in Dong-an, there is a circle shift between the two tones that occur in checked syllables (Cheng 1983). This is significant in part because Dong-an is, genetically speaking, quite distant from the Taiwanese dialects (relative to the size and diversity of the Southern Min dialect group) that display the same kind of

alternation. The whole set of tone sandhi alternations is shown in (15) and a directed graph:

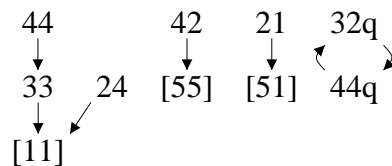
(15) **Dong-an tone sandhi**



In the graph, tones that occur in the base inventory are indicated using Roman numerals I, II, III, and IV (for the four historical tones) and lower-case letters a and b (for the high and low register respectively). Pure sandhi tones are indicated with Chao tone numbers in square brackets.

As should be evident from the graph, there is neutralization at one point in the system, where IIIb and IVb merge to become the low sandhi tone [11]. However, this neutralization does not seem to be related to the circle shift between IVa and IVb (the checked-syllable tones). That is to say, there does not seem to be any respect in which this circle shift makes the output tonal inventory less marked than the input inventory. This is no less true if we look at the system in terms of the phonetic values of the tones:

(16) **Dong-an tone sandhi**



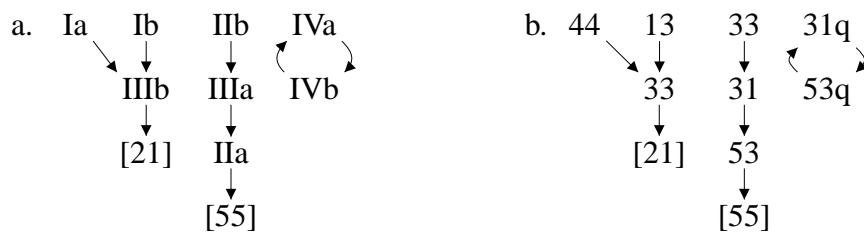
If we look at it in isolation, the neutralization to [11] seems to be part of a sort of lowering chain, and it also allows for the elimination of the rising tone [24]. One possible analysis would be to argue that 44q becomes 32q for the same reasons that 44 becomes 33, presumably a markedness constraint that penalizes 44 (perhaps a constraint against high tones). This raises two questions: first, if such a constraint exists, why does 42 become [55] rather than a lower tone (e.g. [11])? Closer to the point, how can we

motivate the return mapping of 32q to 44q in the face of such a constraint? This is not a trivial problem. Ultimately, it will become a mute point as we look at three cases of circular chain shifts with no neutralization anywhere in the system, namely §4.1.2.4, §4.1.4, and §4.1.4. However, before doing this, we will briefly examine three more *rusheng* circles.

4.1.2.2 Tainan and other Taiwanese dialects

While Taiwanese tone sandhi is closely associated with the single pattern similar to that found in Xiamen (from a dialect we have called Mainstream Taiwanese), Taiwanese actually displays considerable internal diversity and different dialects of Taiwanese have tone sandhi patterns that are very different from the famous circle pattern. One such dialect is that spoken in Tainan (Cheng 1983):

(17) **Tainan tone sandhi**

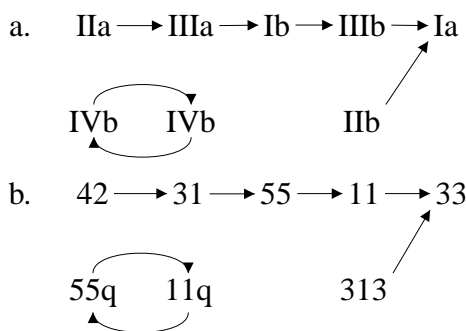


It is similar, in several respects, to the Dong-an system. However, it is quite clear on historical grounds that it has to have developed separately. The other dialects in the group to which Dong-an belongs do not have this type of alternation between *rusheng* tones, but rather have an alternation that looks like a precursor to this ring. The same can be said of Tainan and the similar dialects of Taiwan.

4.1.2.3 Chaoyang

Chaoyang dialect tone sandhi is notable in that it has actually been subject to detailed phonological analysis. Both Yip (2002) and Hsieh (2004) examine this case in terms of contrast preservation/maximization theories. However, it is notable that neither analysis seeks to account for what happens in checked syllables, where there is a circle just as in the dialects discussed above. The whole system may be represented as follows, based on data from Ballard (1988):

(18) **Chaoyang tone sandhi**

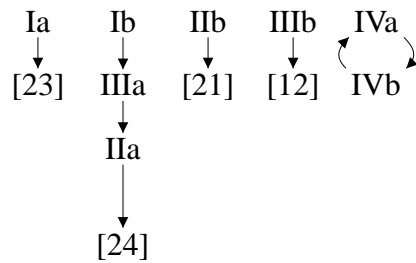


Again, it seems difficult to derive the circle shift from principles of contrast preservation.

4.1.2.4 Chao-an

The most interesting case of a *rusheng* circle that will be discussed here, however, is that found in Chao-an (Ballard 1988). It is interesting precisely because the tone sandhi system is *entirely* free of neutralization, so there is no way of “leveraging” neutralization elsewhere in the system to motivate the circle. The whole set of alternations, in terms of historical categories, is as follows:

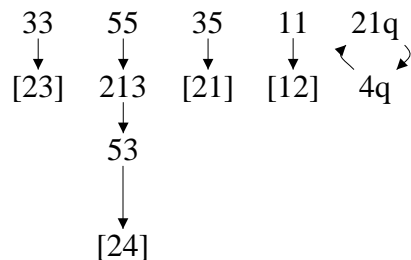
(19) **Chao-an tone sandhi**



If we think about this system as a graph, then no node has more than one incoming edge or more than one outgoing edge—every change is contrast-preserving.

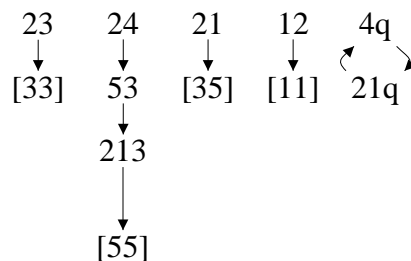
The pattern is no less problematic when it is viewed in terms of phonetic values. As can be seen below, it is hard to argue that the output inventory is less marked than the input inventory:

(20) **Chao-an tone sandhi**



In the output, all three level tones—which should be relatively “unmarked”—are eliminated. While one rising tone becomes a falling tone, a high level tone becomes a falling-rising tone, the mid level tone and high falling tone both become rising tones. We can improve the situation somewhat if we reverse our assumptions about which tones are underlying and which are derived:

(21) **Chao-an tone sandhi**



Under this scenario, two of the rising tones, 23 and 12, are eliminated from the output inventory. However, another rising tone, [35], is added to the inventory. It is not clear, then, that markedness reduction really plays any role in this system. What is even more clear is that there is no sense in which markedness reduction can be said to drive the *rusheng* circle.

4.1.3 Chaozhou (Chaoshan)

Chao-an is not the only Min dialect with both a neutralization-free tone sandhi system and a tone sandhi circle. One such case, historically independent from the other circles and not involving *rusheng*, is that of Chaozhou (Li 1994). The base tones of Chaozhou are given in (22):

(22) **Chaozhou tones (by historical category)** (Li 1994:304)

	<i>ping</i>	<i>shang</i>	<i>qu</i>	<i>ru</i>
<i>yin</i>	┆ (33)	∨ (42)	∩ (213)	┆ (2)
<i>yang</i>	┆ (55)	┆ (35)	┆ (11)	┆ (5)

As in other Southern Min dialects, Chaozhou tones each have a sandhi tone. These are listed below:

(23) **Chaozhou tones and their sandhi tones** (Li 1994:309)

category	<i>Ia</i>	<i>IIa</i>	<i>IIIa</i>	<i>IVa</i>	<i>Ib</i>	<i>IIb</i>	<i>IIIb</i>	<i>IVb</i>
<i>base</i>	┆ (33)	∨ (42)	∩ (213)	┆ (2)	┆ (55)	┆ (35)	┆ (11)	┆ (5)
<i>sandhi</i>	┆ (23)	┆ (24)	┆ (55)	┆/┆ (4/24)	∩ (213)	∩ (21)	∩ (12)	┆/┆ (2/5)

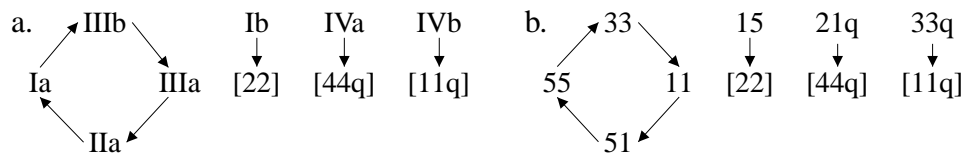
It should be noted that the sandhi tone for IIIa is identical to the base tone for Ib and that, likewise, the sandhi tone for Ib is identical to the base tone for IIIa. For each of the other tone categories, the sandhi tone is a pure sandhi tone; that is, it does not correspond to any of the base tones. This circle shift is important for two reasons. First,

it has clearly developed independently of the tone circles we have examined up to this point. In this circle, IIIa becomes Ib and Ib becomes IIIa, a pattern that is not found in any of the other languages we have examined. Secondly, it reinforces what we have already learned from the Chao-an example, namely that circle shifts need not be driven by neutralization.

4.1.4 Yilan: a circle shift sans neutralization

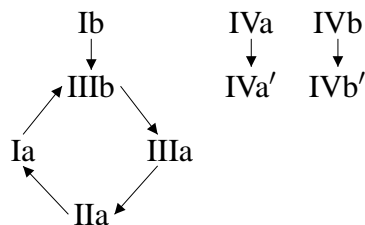
This same point is to be made even more dramatically with the case of Yilan. It will be noted, based upon the following diagram, that the Yilan sandhi system is nearly the same as that of Mainstream Taiwanese (Cheng 1983):

(24) Yilan tone sandhi



Xiamen/Mainstream Taiwanese, for comparison, displays the following system (stated in terms of historical categories, and considering a subdialect without the *rusheng* circle):

(25) The Xiamen/Mainstream Taiwanese tone sandhi circle



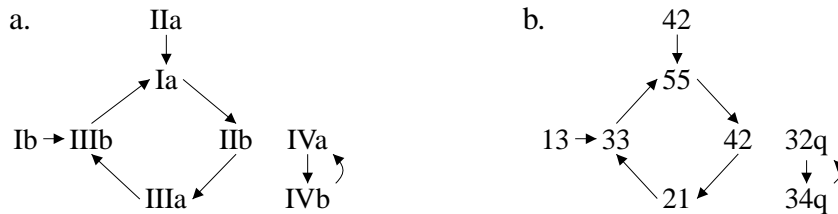
Only a casual inspection is necessary to establish that these have to be descendants of the same system, and that they are identical at some level. What is different, however, is that Ib is mapped in sandhi to a pure sandhi tone rather than to IIIb. A comparison

of the Southern Min tone sandhi systems inventoried by Cheng (1983) and Ballard (1988) show that the Ib → IIIb mapping is one of the oldest and most widespread of the structure preserving tonal alternations in this group. It therefore seems likely that Ib actually “broke away” from IIIb and the rest of the circle at a relatively recent time. The result is a system that is entirely free of neutralization, which cannot therefore be acting as a driving force behind the circle shift.

4.1.5 Dongshan

It should be fairly clear, at this point, that circle shifts do not need to involve neutralization. However, this is not the only problem that Southern Min tone sandhi rings present to contrast preservation theories. Another prediction of Łubowicz (2003) PC Theory is that chain shift scenarios will never involve neutralization at more than one point on the chain. However, the case of Dongshan shows that this prediction is probably incorrect. Consider the following sandhi system:

(26) **Dongshan tone sandhi**



What is perhaps most salient about this pattern is the fact that there are two points of neutralization: both Ib and IIIa become IIIb; both IIIb and IIa become Ia. This makes little sense from a contrast preservation point of view, even if both Ib and IIa are highly marked.

Structural Optimality, by contrast, can generate this type of pattern (as well as those described above) very easily. It is not necessary to go through each of these analyses

individually—it should already be evident from Chapters 2 and 3 how such analyses would work. This will be made more clear, however, by an analysis of the famous Xiamen/ Mainstream Taiwanese in §4.1.5.1.

Before moving on to that analysis, it is interesting to note that the Dongshan circle shift has to have developed separately from that in Xiamen, Mainstream Taiwanese, Yilan, and so forth. In all of those other languages, I**b** and III**b** have merged, and this has produced part of the circle. In Dongshan, however, they remain distinct. The two patterns only share two mappings: I**b** → III**b** and II**a** → I**a**. Since neither of these are part of the circle proper in Dongshan, it is quite clear that the circle developed independently. This is significant in light of the claim that the tone circle is a freakish sort of phenomenon that developed only one time under a special conspiracy of circumstances (Moreton 2004b). In fact, given the right prerequisites, it does not appear that developing a circular chain shift is particularly unusual. In light of this fact, it is reasonable that the theoretical tools necessary to describe and explain such phenomena be developed. We will now show that the framework developed in this dissertation is well-suited to those ends.

4.1.5.1 A Structural Optimality account

The logic behind a Structural Optimality account of circle shifts is quite simple. There is a constraint, HIGHER, which mandates that outputs be higher on some scale than their inputs—it “pushes” mappings up the scale just as it would in a normal chain shift. Because there is an ENDMOST constraint ranked below HIGHER, the upward movements are minimal. This would go on forever if not for the fact that scales are finite. Once it is impossible for an output to be any higher on a scale than the input, it becomes impossible to satisfy ENDMOST. Since ENDMOST dominates SAME (in this case) the next best thing to being higher on the scale is to be at the bottom of the scale. Thus, after climbing incrementally to the top of the scale, we drop abruptly to

the bottom.

This explains how the circle itself works, and is basically adequate for cases like Yilan. However, many circles have what we might call “danglers”—elements that are mapped to an element in the circle but which have no element mapped to them. These could actually represent two different possibilities within Structural Optimality. The theoretically more “pure” claim is that there is a second scale and that the danglers sit high on this scale while the elements in the chain sit at the bottom end. Because of an ENDMOST constraint, there is an incentive for the danglers to drop to the level of the chain elements but no incentive for the chain elements to rise to the level of the danglers. The other solution, which will be adopted in the analysis given below, is to posit that the danglers are so highly marked that it is cheaper to skip the step on the scale at which they reside than to allow an input to be mapped to them. That is to say, they are affected by markedness constraints that dominate ENDMOST. In a sense, this approach may capture insights similar to those captured by Barrie’s (2006) PC Theory analysis. The difference is that it does not make the incorrect prediction about the number of points of neutralization than can be present in a chain-shift scenario.

The most problematic part of an analysis in these terms is actually deciding where to divide the ring to form a single scale. This is difficult not because it presents any intricate difficulties but precisely because the different solutions are so difficult to distinguish. This is not to say that they are always indistinguishable, but only that this fact is under-determined in most cases. Somewhat arbitrarily, we will employ the following scale:

$$(27) \quad T = \{24\} < \{22\} < \{21, 32q\} < \{53, 54q\} < \{44\}$$

This places the rising tone 24, the “dangler,” at the bottom of the scale and the high level tone 44 at the top of the scale. The checked tones (*rusheng*), indicated, as elsewhere, with a -q, are placed on the scale with smooth-syllable tones to which they are pho-

netically similar (and with which they may be underlyingly identical). We will assume that there is a high-ranked faithfulness constraint (or, more likely, a high-ranked set of faithfulness constraints) that prevent checked syllables from being mapped to smooth syllables and thus limit the possible outputs for underlying checked tones to the two tones 32q and 54q.

The remainder of the rankings are identical to (one of) those developed in 2.7.5 with one exception: there is a constraint *RISE ranked below HIGHER[T] but above ENDMOST[T] which penalizes the occurrence of the rising tone 24 in outputs. Thus, as shown in (28), 24 will be mapped to the next point on the scale, namely, 22:

(28) **Tableau for /24/**

	24	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24		*!	*		
(b)	22				*	*
(c)	21				**!	*
(d)	53				**!*	*
(e)	44				**!**	*
(f)	32q	*!			**	*
(g)	54q	*!			***	*

However, as can be seen in 29, when it is not possible to go any higher on the scale because the input is 44, and therefore at the top, the output does not “fall” all of the way to 24 because the markedness constraint against rising tones dominates the ENDMOST constraint, which penalizes outputs which are not at the bottom of the scale.

(29) **Tableau for /44/**

	44	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24		*	*!		*
(b)	22		*		*	*
(c)	21		*!		**	*
(d)	53		*!		***	*
(e)	44		*!		****	*
(f)	32q	*!	*		**	*
(g)	54q	*!	*		***	*

The rest of the mappings for the smooth syllables are like those for normal chain shifts: the output will be the option that best satisfies ENDMOST[T], by being maximally close to the bottom of the scale, while still satisfying HIGHER[T]:

(30) **Tableau for /22/**

	22	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24		*!	*		*
(b)	22		*!		*	
(c)	21				**	*
(d)	53				***!	*
(e)	44				****!	*
(f)	32q	*!			**	*
(g)	54q	*!			***	*

(31) **Tableau for /21/**

	21	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24		*!	*		*
(b)	22		*!		*	*
(c)	21		*!		**	
☞ (d)	53				***	*
(e)	44				****!	*
(f)	32q	*!	*		**	
(g)	54q	*!			***	*

(32) **Tableau for /53/**

	53	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24		*!	*		*
(b)	22		*!		*	*
(c)	21		*!		**	*
(d)	53		*!		***	
☞ (e)	44				****	*
(f)	32q	*!			**	*
(g)	54q	*!			***	

For checked syllables, the situation is related but slightly different. Because FAITHCHK is ranked so highly, the scale is effectively shortened relative to its interaction with HIGHER[T] and ENDMOST[T]. For practical purposes, there are only two outputs that need to be considered: when the input is 32q, it is possible to satisfy HIGHER[T], and even though this incurs one more violation of ENDMOST[T] than the identity candidate would have, 54q is the best candidate.

(33) **Tableau for /32q/**

	32q	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24	*!	*	*		*
(b)	22	*!	*		*	*
(c)	21	*!	*		**	
(d)	53	*!			***	*
(e)	44	*!			****	*
(f)	32q		*!		**	
(g)	54q				***	*

However, if the input is 54q, satisfying HIGHER[T] without violating FAITHCHK is impossible. It is possible, to improve on the input with regards to “markedness,” specifically ENDMOST[T], so the optimal output is 32q:

(34) **Tableau for /54q/**

	54q	FAITHCHK	HIGHER[T]	*RISE	END[T]	SAME[T]
(a)	24	*!	*	*		*
(b)	22	*!	*		*	*
(c)	21	*!	*		**	*
(d)	53	*!	*		***	
(e)	44	*!			****	*
(f)	32q		*		**	*
(g)	54q		*		***!	*

A minimal circle shift with no neutralization in checked syllables is the result of this ranking.

It should be clear based on this demonstration and the earlier explorations of the type of mapping scenarios that can be produced in Structural Optimality, that the other patterns that we have presented above yield as easily to the same kind of analysis. In other words, there can be little question whether Structural Optimality is powerful enough as a theory to capture these patterns. The only question is whether there is some

less-powerful theory that can capture all of the attested patterns with more insight. This is a somewhat open-ended question that cannot be answered at this early point, when these phenomena are still so poorly understood. However, it can be said that contrast preservation theories on the order of Łubowicz (2003), Barrie (2006), and Hsieh (2004), are not the key to understanding circle shifts in Southern Min tone systems. This will be the subject of §4.1.6.

4.1.5.2 Relevance for contrast-preservation theories

We have reviewed several cases, a number of them independent developments, that are incompatible with contrast-preservation theories of chain shifts. It is not simply the case that these patterns cannot be explained insightfully in terms of contrast preservation. Rather, they run counter to the fundamental predictions of theories of this kind.

The first set of cases that we examined were circle shifts with no neutralization. As was mentioned briefly before, PC Theory holds that chain shifts are driven by markedness reduction: there is some very highly-marked structure, at the “beginning” of the chain, and the fundamental impetus for the shift is the reduction of the markedness of the output inventory by mapping this highly-marked structure to some less-marked structure. In other words, all chains are push chains (rather than drag chains). Under rankings that produce chain shifts, the imperative to avoid neutralizations is higher than the imperative to avoid unfaithful input-output mappings, so structures that are underlyingly identical to the less-marked structure to which the first structure is mapped in the output are mapped to the next-best output (in terms of markedness and faithfulness). This same pattern continues as a kind of chain reaction until there is neutralization at the end of the chain or a “non-structure preserving” mapping of an input to a structure that does not otherwise appear in the output. The most basic prediction of this type of theory is that the output inventory in a chain shift scenario will always be less marked than the input inventory. For a circle shift, this means that there must be neutralization

since the input inventory would otherwise be identical to the output inventory. However, we have seen that circle shifts without neutralization not only exist but are also quite common relative to the frequency of circle shifts of any kind. This is the one reason for rejecting PC Theory of circle shifts in favor of a theory including anti-identity relations like Structural Optimality and the theory of logical scales.

As has been mentioned, though, there is another—more subtle—prediction that is made by PC Theory which is also contradicted by Southern Min tone circle data. It appears that Łubowicz’s (2003) PC Theory predicts that chain shifts where there is neutralizations at multiple points on in a chain shift scenario should exist. However, as we have shown with the Dongshan example in §4.1.5 such cases can exist in actual languages. This too seems incompatible with the central notion of contrast preservation theory.

4.1.6 The historical evolution of circle shifts in Southern Min

The discussion so far has shown how circle shifts can be modeled in the grammar. It has not explained, however, why they should actually exist and why they should be largely (perhaps entirely) confined to tone systems. I will now show how these patterns likely came into being, starting with a general theoretical discussion about the development of patterns of alternation in complex tone systems and how these developments may be formalized. I will then proceed to specific analyses of the development of the Xiamen and Dongshan circles as well as the *rusheng* rings.

4.1.6.1 First principles

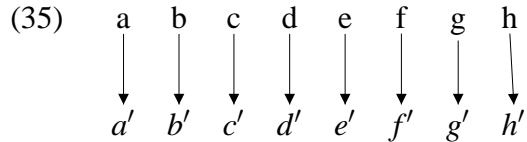
One useful way of understanding a tone sandhi system is to view it as a directed graph (digraph) where nodes represent tones (in the Asianist sense) and edges (arrows) represent mapping relationships between inputs and outputs in some context. We have already seen diagrams that reflect this view of tone sandhi relationships in §4.1.1–4.1.5.

Such graphs must have the special property that each node has, at most, one outgoing edge. Such a graph (indeed, all graphs) are, at some level of abstraction, identical to relations (consisting of a set of ordered pairs). This way of looking at tone sandhi is not new, and is found in earlier works like Court (1985). Here, I will attempt to give a more rigorous formulation of this model.

Historical changes in a system like this can take a number of different forms. Two nodes can be “merged” representing a merger in the tonal system; in these cases (informally speaking), all of the incoming edges for the old nodes become incoming edges for the new node. Thus, if there were edges (a, b) and (c, d) in a graph and b and d merged as e , the new graph would include the edges (a, e) and (c, e) . Tonal splits, by the same token, are represented by the replacement of one node with two nodes; under this condition, if there was an outgoing edge from the old node leading to some other node, there will be edges from the two new nodes to the other node. In other words, if there is an edge (a, b) and a splits to become a and a' , the graph will then include the edges (a, b) and (a', b) . Interesting problems result when a node with an incoming node splits. Since, by definition, the source node can have at most one outgoing node, it cannot be mapped to both of the new nodes, so something external to this graph theoretic model must decide which of the new nodes becomes the destination of the edge from the source-node. The same issue arises when there is a merger between nodes with outgoing edges pointing to different nodes.

What leads to mergers and splits in the first place? The same principles that motivate mergers and splits in other phonological domains. However, in the case of mergers, complex tone systems of the Southern Min type have a special property, namely, that the phonetic values of the tones are much less stable, historically speaking, than the categories to which they belong. This means that “allotones” of the same may quite easily become quite different from one another and the range of possible mergers be-

comes quite high. If we imagine a massive split, where all of the tones come to have two allotones—and this appears to have been the case in earlier Southern Min—we also imagine a system that is ripe for the development of complex tonal patterns like circle shifts. We could represent a system like that as follows, where plain letters represent base tones and letters with primes represent pure sandhi tones):



We will assume that this is identical to the initial state of the Min tone sandhi system.

There is one other process that we must consider, which we might call dissociation. In this process, a tone is “detached” from its sandhi tone and is (in most cases) mapped to a different output, often a pure sandhi tone. This is one of the strongest pieces of evidence that the alternations we are examining here are not simply the artifact of historical changes (in the conventional sense, with no synchronic component). Were it true that these patterns were simply the result of splits and mergers among the tones of pairs of allomorphs, it is not easy to understand why a change of this type would take place, since it would involve the reversal of a merger, and thus an exception to one of the most fundamental axioms of sound change. Instead, it must be the case that there is an underlying tone for both allomorphs and that the phonological grammar generates the surface allomorphs.

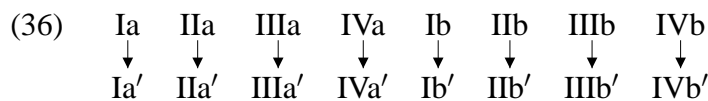
At this point, it is also important to mention a curious but significant fact: pure sandhi tones are far more likely to merge with either sandhi tones or base tones than base tones are to merge with one another. There are various reasons that this could be the case, but there is not space here to give a complete discussion of these possibilities. For our purposes, it is sufficient simply to mention that this tendency exists, since it will be evident in the discussion below.

We will now apply these principles to understanding the development of the tone

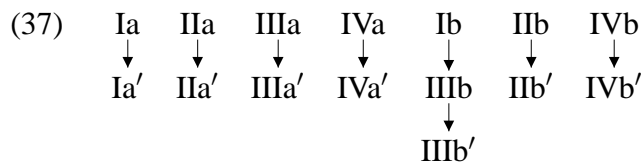
circles in Xiamen and Dongshan, with an eye to showing how these two circle shifts must have developed independently. In these expositions, and in some cases about, I will refer to tones according to their Middle Chinese categories using Ia, Ib, IIa, IIb, and so forth.

4.1.6.2 Xiamen and Mainstream Taiwanese

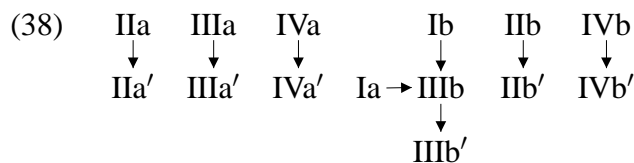
In reconstructing the tone sandhi system of Xiamen/Mainstream Taiwanese, we will begin with the assumption that, at the earliest stage, Southern Min tone sandhi involved eight base tones with eight distinct sandhi tones. The later patterns are primarily the result of mergers of the eight tones and (especially) their sandhi allotones.



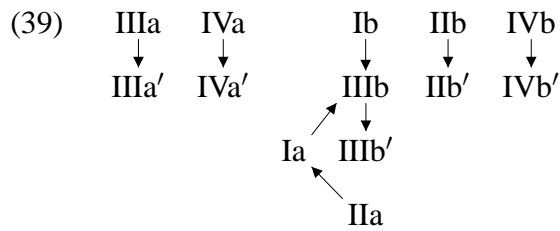
The first change that occurred in this system, based on comparative evidence, was the merger of Ib' and IIIb. The merger of these two tones is attested very widely in Southern Min, far outside of the subgroup to which Xiamen belongs. This resulted in a system containing a chain shift but no neutralization:



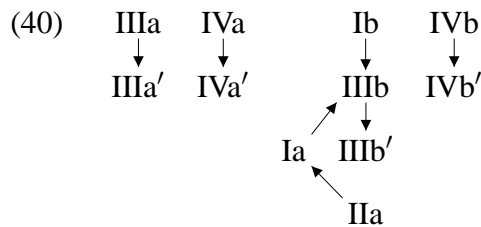
Then, another widely attested innovation occurred: the merger of Ia' with IIIb, meaning that the distinction between Ia and Ib was neutralized in sandhi context:



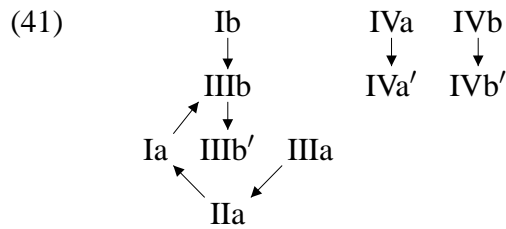
One of the chain shifts was subsequently lengthened by the merger of IIa' with Ia:



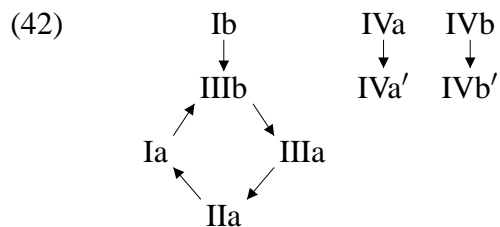
It was after this innovation that Dongshan dialect, as described in §4.1.6.3, diverged from the dialect group that includes Xiamen and Mainstream Taiwanese. The next change, a merger of IIb and IIIb, is shared by the languages in this group, but not by Dongshan. The resulting system can be depicted as follows:



Next, the chain shift was lengthened still further by the merger of IIIa' with IIa:



The final step, leading to the creation of the grand tone circle, was the merger of IIIb' and IIIa, completing the circle:



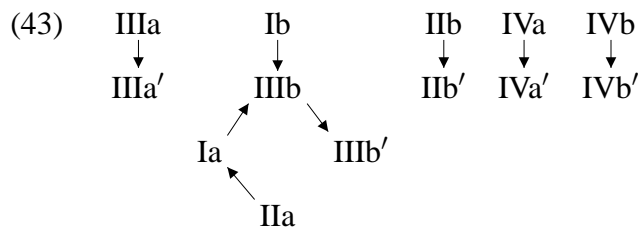
In many of the other systems, IVa and IVb form a circle shift, as described above in §4.1.2. On comparative evidence from various Xiamen dialects, it appears that IVa' first merged with IVb, then IVb' merged with IVa. The fact that this pattern occurs

sporadically in dialects that are not necessarily closely allied to one another suggests that a development of this type occurred a number of times independently.

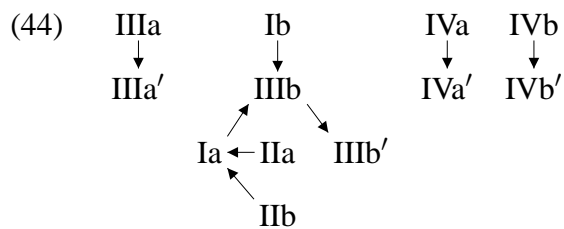
4.1.6.3 Dongshan

A different course of events has to be invoked in order to explain the tone sandhi pattern in Dongshan, already described above in §4.1.5. The fact that the course of development is different from that for Xiamen, Mainstream Taiwanese, Longxi, and Gaoxiang, is very significant: the grand kind of tone circle seen in Xiamen has developed more than once—more or less independently.

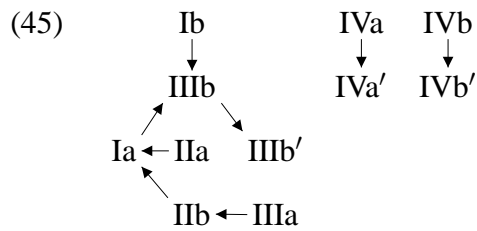
At the point of divergence between the two dialect groups, the tone sandhi system was as follows:



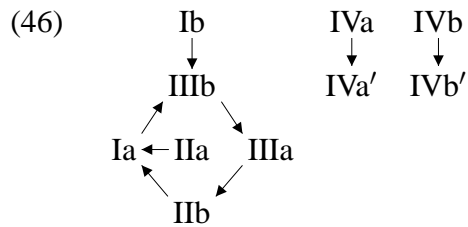
However, there was no merger between IIb and IIIb. Instead, IIb' merged with Ia.



Then two other changes occurred which closed the circle. The relative order of these two events cannot be established with any degree of certainty, but one of the two possible orders will be chosen here for purposes of exposition. First, IIIa' merged with IIb:



The the circle was closed when IIIb' merged with IIIa:



What this demonstrates is that the Dongshan circle actually came to be through a set of historical innovations that are distinct from those that created the Xiamen/Mainstream Taiwanese circle. In other words, it appears the circular chain shifts exist in Southern Min languages not because they came into being once through a set of processes that could never recur in another language family, but because the tone system and tone sandhi system of Southern Min languages are prone to developments of complex patterns of alternation that include both chain shifts of unusual length and circular chain shifts. Next we will see further evidence that tonal circle shifts are not confined to Southern Min alone, but are to be found in other (distantly related or unrelated) languages including the Hmong-Mien language A-Hmao and the Tibeto-Burman language Jingpho.

4.2 A-Hmao Bounce-back Effects

We have already talked briefly about tone sandhi in the Far-Western Hmongic language A-Hmao. In the discussion in § 3.2.2, we examined a tonal chain shift that affected tones in the “upper” tonal register. In Western A-Hmao, for example $\uparrow H \uparrow$ is mapped to H in sandhi context (after HM and LM) but H is mapped to M in the same environment.

Examples of this pattern are given below (note that the tone letters 1, 1, 1, 1, and 1 represent ↑H, H, M, HM, and LM respectively):

- (47) a. ki1 ‘road, way’ kau\ki1 ‘custom’
 b. mpa1 ‘pig’ dfi|o\mpa1 ‘lard’

4.2.1 Western A-Hmao

On closer examination, this does not exhaust the set of tone sandhi alternations that occur in the various dialects of A-Hmao. In the lower tonal registers of both Eastern and Western A-Hmao, there are “bounce-back” effects—patterns where all other inputs are mapped to a single output, but the input identical to that output is mapped to the “least-marked” *non-identical* output. This idea was discussed in § 2.7.3. Here, we will deal with it more concretely and more completely.

In the “lower” tonal register of Western A-Hmao, there are four tones: LM, M, L, and ↓L (represented in transcriptions as 1, 1, 1 and 1). We will be concerned with M, L, and ↓L at this point because LM patterns with HM (the other tone that triggers sandhi) rather than the rest of the lower register tones. In sandhi environment, both L and ↓L become M; however, M becomes L. These patterns are illustrated in the data below:

- (48) a. dzfi1 ‘cold’ ?au\dzfi1 ‘cold water’
 b. nfi1 ‘rain’ ?au\nfi1 ‘rain water’
 c. ndfil1 ‘glutinous’ ndfili\ndfil1 ‘glutinous rice’

A set of ranking schemas that generate neutralization with bounce-back were described in Chapter 2. There, it was shown that neutralization with bounce-back is generated when DIFF dominates SAME, ENDMOST, and HIGHER and where ENDMOST dominates HIGHER. Since the ranking of SAME is not particularly crucial (as long as it is ranked below DIFF) it will not be included in tableaux or further discussion². With this

²It is not difficult to show that SAME is always vacuous when it is ranked below DIFF. This is due to the fact that they have complementary violation profiles.

theoretical foundation, we may propose the following scale and constraint ranking:

- (49) a. $T = \{M\} < \{L\} < \{\downarrow L\}$
 b. $\text{DIFF}[T] \gg \text{ENDMOST}[T] \gg \text{HIGHER}[T]$

If we combine this with the scale that we had already developed for the Western Hmong high register shift (translated, of course, into Western A-Hmao terms) with this scale, we get the following:

- (50) $T = \{M', \uparrow H\} < \{L, H\} < \{\downarrow L, M\}$

On the one hand, it is surprising that the two M tones, M' and M , should be at opposite ends of the scale. On the other hand, it is comforting to note that the scale is essentially a parallel list of the tones from each register in descending pitch. There is another issue to be addressed, which is the register distinction. This too can be encoded as a scale, which we will call R :

- (51) $R = \{\downarrow L, L, M'\} < \{M, H, \uparrow H\}$

The primary importance of R , at this point, lies in the generalization that tone sandhi processes in A-Hmao and other Western Hmongic languages do not involve alternations in register³. In Chapter 5, it will be seen that register scales function in other ways in the phonologies of Western Hmongic languages, but for our purposes we need only one constraint, $\text{SAME}[R]$, which may safely dominate all of the other constraints. At this point, the ranking of $\text{SAME}[R]$ is non-crucial: there is no benefit to changing register, so even a very low-ranked $\text{SAME}[R]$ would eliminate candidates that do so.

Taking these constraint rankings and scales, we predict the proper outputs for all of the low-register inputs:

³There is actually an exception to this in A-Hmao in a tone sandhi process that we will not consider in depth in this work. This should not be a matter of stress, however—in an Optimality Theory world, exceptions to generalizations are to be welcomed.

(52) a. **Tableau for $\downarrow L$ (correct output)**

	$\downarrow L$	SAME[R]	DIFF[T]	END[T]	HIGHER[T]
(a)	$\downarrow L$		*!	**	*
(b)	L			*!	*
(c)	M'				*
(d)	M	*!	*		*
(e)	H	*!		*	*
(f)	$\uparrow H$	*!		**	*

b. **Tableau for L (correct output)**

	L	SAME[R]	DIFF[T]	END[T]	HIGHER[T]
(a)	$\downarrow L$			*!*	
(b)	L		*!	*	*
(c)	M'				*
(d)	M	*!		**	
(e)	H	*!	*	*	*
(f)	$\uparrow H$	*!		**	*

c. **Tableau for M' (correct output)**

	M'	SAME[R]	DIFF[T]	END[T]	HIGHER[T]
(a)	↓L			**!	
(b)	L			*	
(c)	M'		*!		*
(d)	M	*!		**	
(e)	H	*!		*	
(f)	↑H	*!	*	**	*

This is to be expected, since this result had already been discovered in abstract terms and this is simply a concretization of that preexisting finding. The issue of capturing both the upper-register alternations and the lower-register alternations with the same grammar is somewhat more complicated. In the higher register, $\uparrow H \rightarrow H$ and $H \rightarrow M$ (in sandhi context); M remains M . We may depict this succinctly as follows:

(53) **Upper-register tone sandhi in Western A-Hmao**

$$\uparrow H \rightarrow H \rightarrow M$$

If we evaluate the same set of candidates paired with the upper-register tones as inputs, the result is not what is desired:

(54) a. **Tableau for M (wrong output)**

	M	SAME[R]	DIFF[T]	END[T]	HIGHER[T]
(a)	↓L	*!	*	**	*
(b)	L	*!		*	*
(c)	M'	*!			*
(d)	M		*!	**	*
(e)	H			*	*
(f)	↑H			**!	*

b. **Tableau for H (wrong output)**

	H	SAME[R]	DIFF[T]	END[T]	HIGHER[T]
(a)	↓L	*!		**	
(b)	L	*!	*	*	*
(c)	M'	*!			*
☹ (d)	M			**	
(e)	H		*!		
(f)	↑H			**	*!

c. **Tableau for ↑H (correct output)**

	↑H	SAME[R]	DIFF[T]	END[T]	HIGHER[T]
(a)	↓L	*!		**	
(b)	L	*!		*	
(c)	M'	*!	*		*
(d)	M			**!	
☹ (e)	H			*	
(f)	↑H		*!	**	*

One of the mappings is correct, but this seems to be fortuitous. A cursory examination of the mappings reveals that a reversal of the scale would not yield the correct outputs either. Under that condition, ↑H and H would both become M, but M would become H. This is trivially true, since tonal behavior in the upper register should be the same as that in the lower register (topologically speaking) and it was demonstrated in Chapter 2 that this constraint ranking will always produce mappings of this type. It is absolutely necessary, then, that there be some mechanism that distinguishes between registers. Here we will employ variants of the scale-referring constraints that we have proposed which are directly sensitive to positions in *R*; that is, only penalize tones that are at the

top of R . It is worthy of note that such constraints could easily be constructed using local constraint conjunction, though one would be wary to admit the type of power that constraint conjunction would give to a model that is admittedly very powerful already. Be that as it may, we will employ such register-sensitive constraints without giving a fully theoretical justification for their existence. The specific constraints that we will need are $\text{SAME}[T]\uparrow$ and $\text{HIGHER}[T]\uparrow$. These constraints are just like $\text{SAME}[T]$ and $\text{HIGHER}[T]$ except that they can only be violated *when the output tone is in the upper register*.

The ranking logic for these constraints is simple, and was already addressed in § 2.7.4 and § 3.2. $\text{HIGHER}[T]\uparrow$ must dominate $\text{SAME}[T]\uparrow$ and $\text{SAME}[T]\uparrow$ must dominate $\text{END}[T]$ and $\text{DIFF}[T]$. $\text{SAME}[R]$ must dominate all of these constraints. Otherwise, candidates could satisfy the upper-register constraints just by being in the lower register (with no change along T). When this logic is tested, it proves to be correct:

(55) **Tableau for $\downarrow\text{L}$ (correct output)**

	$\downarrow\text{L}$	$\text{SAME}[R]$	$\text{HIGHER}[T]\uparrow$	$\text{SAME}[T]\uparrow$	$\text{DIFF}[T]$	$\text{END}[T]$
(a)	$\downarrow\text{L}$				*!	**
(b)	L					*!
(c)	M'					
(d)	M	*!	*		*	
(e)	H	*!	*	*		*
(f)	$\uparrow\text{H}$	*!	*	*		**

(56) **Tableau for L (correct output)**

L	SAME[R]	HIGHER[T]↑	SAME[T]↑	DIFF[T]	END[T]
(a) ↓L					*!*
(b) L				*!	*
(c) M'					
(d) M	*!		*		**
(e) H	*!	*		*	*
(f) ↑H	*!	*	*		**

(57) **Tableau for M' (correct output)**

M'	SAME[R]	HIGHER[T]↑	SAME[T]↑	DIFF[T]	END[T]
(a) ↓L					**!
(b) L					*
(c) M'				*!	
(d) M	*!		*		**
(e) H	*!		*		*
(f) ↑H	*!	*		*	**

(58) **Tableau for M (correct output)**

M	SAME[R]	HIGHER[T]↑	SAME[T]↑	DIFF[T]	END[T]
(a) ↓L	*!			*	**
(b) L	*!				*
(c) M'	*!				
(d) M		*		*	**
(e) H		*	*!		*
(f) ↑H		*!	*		**

(59) **Tableau for H (correct output)**

H	SAME[R]	HIGHER[T]↑	SAME[T]↑	DIFF[T]	END[T]
(a) ↓L	*!				**
(b) L	*!			*	*
(c) M'	*!				
(d) M			*		**
(e) H		*!		*	
(f) ↑H		*!	*		**

(60) **Tableau for ↑H (correct output)**

↑H	SAME[R]	HIGHER[T]↑	SAME[T]↑	DIFF[T]	END[T]
(a) ↓L	*!				**
(b) L	*!				*
(c) M'	*!			*	
(d) M			*		**!
(e) H			*		*
(f) ↑H		*!		*	**

This grammar, then, is adequate for the whole set of Western A-Hmao tone sandhi relationships that we have discussed.

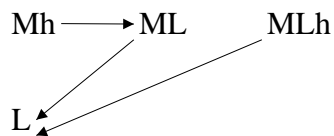
4.2.2 Eastern A-Hmao

The eastern dialect of A-Hmao also has tone sandhi with bounce-back. The whole tone sandhi system of Eastern A-Hmao is incredibly complicated and a full analysis would constitute a lengthy paper by itself. What will be presented here, however, will be an analysis of a significant subset of the tonal alternations in the languages. I will discuss Eastern A-Hmao at this point for two reasons: First, the data illustrating the bounce-back effect are somewhat more voluminous for Eastern A-Hmao than for Western A-Hmao, though the generalizations about tonal alternations have been made by highly-qualified and competent linguists in both cases, and so there is little reason to doubt

that Western A-Hmao tone sandhi acts in the fashion described here. However, seeing more data is always comforting, and Eastern A-Hmao provides the opportunity to do so. It is also useful to see independent evidence for a putative phenomenon, and Eastern A-Hmao provides that. Though the dialects are closely related, and while the historical precursors for the bounce-back effects in the two dialects are identical, the specific historical processes that led to the emergence of bounce-back in the two cases appear to have occurred independently.

Proto-A-Hmao had an eight-way tonal contrast in monosyllables, like any well-behaved Far-Western-Hmongic language. In A-Hmao, however, three of the lower-register tones have undergone a split, as described by Wang and Wang (1986) and further analyzed by Ratliff (1992a). This split occurred only in nouns and is apparently the result of a morphological marker—probably a prefix—that conditioned the change and subsequently disappeared. This expansion of the tonal inventory led to a concomitant increase in the complexity of the tone sandhi system. Prior to this split, the tones sandhi system of Eastern A-Hmao looked much like Western A-Hmao and, indeed, those of other Far-Western Hmongic languages. We can depict the reconstructed complex of tonal alternations in the lower-register as in (61). As elsewhere, arrows point from underlying forms to the surface forms that occur in sandhi context (note that lowercase “h” indicates breathiness):

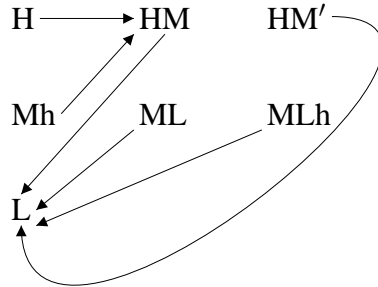
(61) **Proto-A-Hmao tone sandhi**



The Western A-Hmao tone sandhi pattern came about when the L tone merged with Mh (the mid-breathy tone). The Eastern A-Hmao tone pattern underwent a series of changes before its circle shift developed. First, the tonal split discussed above occurred, yielding the system depicted in (62). Mh split into Mh and H, ML split into ML and

HM, and (appropriately enough) MLh split into MLh and HM':

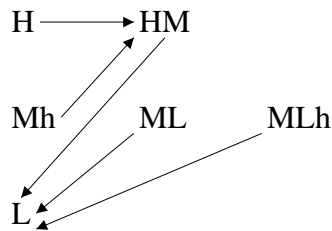
(62) **Pre-A-Hmao tone sandhi: stage 1**



At this point, all of the tones that formerly became L in sandhi context, and all of the tones that split from them, still become L. Unsurprisingly, H, which split from Mh, is mapped to HM, which split from ML. What is more striking—and crucial—is that Mh was also mapped to HM at this point, rather than to ML (which might have been expected).

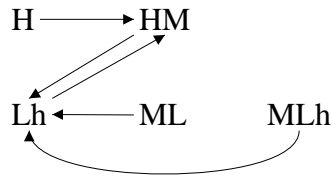
Whether at the time of the nominal/non-nominal tone split or soon thereafter, HM and the tone we have called HM' merged, yielding the simpler system that we see in (63):

(63) **Pre-A-Hmao tone sandhi: stage 2**



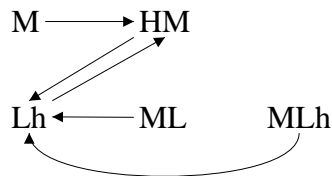
The next innovation was what actually produced the circle shift. Mh and L merged as Lh. All of the tones that became L in sandhi context, and all of the tones that became Mh in sandhi context, were now mapped to this single tone in that environment. This “new” tone, like Mh, became HM in sandhi context, resulting in the exchange shown in (64):

(64) **Pre-A-Hmao tone sandhi: stage 3**



The final change did not affect the structure of the system, but only the gross phonetics of one of the tones: H was lowered to M, replacing Mh after its merger with L:

(65) **Eastern A-Hmao tone sandhi**



We have, in (65), the tone sandhi pattern that currently exists in the Eastern A-Hmao lower register, complete with circle shift and modern tone values.

Even though this circle shift resulted from the merger of proto-Mh with proto-L, just like that in Western A-Hmao, it is clear that they must have occurred separately: Western A-Hmao did not undergo the nominal/non-nominal tone split, while Eastern A-Hmao did, but the Eastern A-Hmao pattern requires that the tone split occurred before the merger of proto-M with proto-L. The resulting pattern is interesting because the bounce-back is apparently being “fed” from two sides—it is difficult to tell what the unmarked endpoint of the scale might be. As we will see, there are actually two scales involved in this case.

Before we undertake the analysis in which this will be demonstrated, it is instructive to look at the data that illustrate these alternations, as taken from Wang and Wang (1986):

(66) **Examples of tone sandhi in Eastern A-Hmao**

	M	→	HM	
a.	zo┐	‘village’	lu┐zo┐	‘the village’ (‘CLF’ + ‘village’)
b.	nu┐	‘horse’	t┐lau┐nu┐	‘horse hair’ (‘hair’ + ‘horse’)
c.	mbə┐	‘fish’	dlfio┐mbə┐	‘fish oil’ (‘oil’ + ‘fish’)
d.	nu┐	‘horse’	ŋyfiəi┐nu┐	‘horse meat’ (‘meat’ + ‘horse’)
	HM	→	Lh	
e.	za┐	‘comb’	lu┐zfiə┐	‘the comb’ (‘CLF’ + ‘comb’)
f.	nau┐	‘rain’	au┐nfiəu┐	‘rain water’ (‘water’ + ‘rain’)
g.	zo┐	‘strength’	dlfiəu┐zfiə┐	‘vigor’ (‘energy’ + ‘strength’)
h.	ŋau┐	‘time’	dzfiəi┐ŋfiəu┐	‘time’ (‘time’ + ‘time’)
	Lh	→	HM	
i.	ŋgfiu┐	‘lazy’	tu┐ŋku┐	‘lazy person’ (‘boy’ + ‘lazy’)
j.	dzfiə┐	‘cool’	?au┐dzie┐	‘cool water’ (‘water’ + ‘cool’)
k.	lfiəu┐	‘old’	ŋfiu┐lfiəu┐	‘old cattle’ (‘cattle’ + ‘old’)
l.	dzfiə┐	‘cool’	ŋgfiəi┐dzfiə┐	‘cool meat’ (‘meat’ + ‘cool’)
	ML	→	Lh	
m.	ba┐	‘hug’	tu┐bfia┐	‘adopted son’ (‘hug’ + ‘boy’)
n.	dlo┐	‘fat’	qai┐dlo┐	‘fat chicken’ (‘chicken’ + ‘fat’)
o.	dlo┐	‘fat’	ŋfiu┐dlo┐	‘fat cattle’ (‘cattle’ + ‘fat’)
p.	da┐	‘die’	zfiəu┐da┐	‘dead sheep’ (‘sheep’ + ‘die’)
	MLh	→	Lh	
q.	ndlfia┐	‘sticky’	ti┐ndlfia┐	‘mud’ (‘earth’ + ‘sticky’)
r.	zfiə┐	‘guilt’	ŋfiy┐zfiə┐	‘sin’ (‘guilt’ + ‘wrong’)
s.	ndlfia┐	‘sticky’	ndlfiəi┐ndlfia┐	‘sticky rice’ (‘rice’ + ‘sticky’)

Looking at these data, and at the diagram in (65), it is easy to see that all of the tones become one of two tones in sandhi context. ML and HM became Lh and the rest of the tones become HM. This leads us to believe that Lh and HM have something in common: it appears that they are at the same point on a certain scale. Since all of the other tones become Lh, it seems likely that Lh and ML are at the end of a scale and there is a neutralization driven by an ENDMOST constraint. However, there is clearly an anti-identity requirement along this same scale because M and Lh are mapped not to themselves, as we would expect if ENDMOST dominated DIFF, but to HM, which we may posit as a member of the next rung of the scale. We should also place ML and MLh at a higher level on the scale, or find some explanation for the fact that Lh and M are mapped to HM rather than one of these tones. Let us call the first scale *T*.

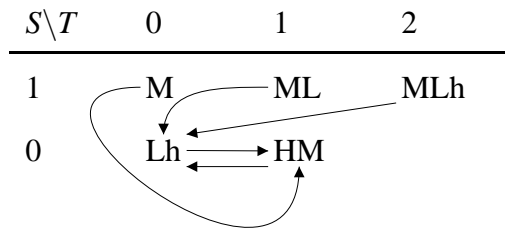
It is also clear that Lh and HM share something in common, since they are the only targets of the neutralization. By logic similar to that we have employed above, we would assume that these two tones are at the bottom of some scale, which we will call *S*. M must be at the top of this scale—otherwise HM, ML, and MLh might be mapped to it under sandhi, rather than to Lh. This leaves us, though, with very little information about the absolute or relative positions of ML and MLh, on either *S* or *T*. The only we know is this: if these tones are at the bottom of *S*, they must be higher than HM on *T*. If they are not at the bottom of *S*, that must be at least as high as *T* on *S*. The nicest assumption, which will be shown to be perfectly workable, is to place them at the same level as *M* on *S*, but at the subsequent levels on *T*. These leaves us with the configuration in (67). Note that the numbers have no theoretical status and are there simply to make it easier for the reader to determine the direction on the table in which the scale progresses:

(67) **Eastern A-Hmao tone scales**

$S \backslash T$	0	1	2
1	M	ML	MLh
0	Lh	HM	

If we add arrows for the tone sandhi alternations, we get the following diagram (68), which is actually far simpler than it looks:

(68) **Eastern A-Hmao tone sandhi by scales**



To state the pattern in intuitive terms, ML, MLh, and HM are all neutralized to Lh because it is judged to be the best output by $ENDMOST[T]$ and $ENDMOST[S]$. However, there is a constraint $DIFF[T]$ that dominates $ENDMOST[T]$. Because of this Lh cannot remain Lh (even though the $ENDMOST$ constraints favor it) and M cannot become Lh either. The next best tones according to $ENDMOST[T]$ are ML and HM. Because $ENDMOST[S]$ is undominated by other constraints on S , HM is favored above ML, so both M and Lh are mapped to HM. This is easily demonstrated with tableaux:

(69) a. **Tableau for ML**

	ML	$END[S]$	$DIFF[T]$	$END[T]$	$SAME[S]$	$SAME[T]$
(a)	M	*!				*
(b)	ML	*!	*	*		
(c)	MLh	*!		**		*
(d)	Lh				*	*
(e)	HM		*!	*	*	

b. **Tableau for M**

	M	END[S]	DIFF[T]	END[T]	SAME[S]	SAME[T]
(a)	M	*!	*			
(b)	ML	*!		*		*
(c)	MLh	*!		**		*
(d)	Lh		*!		*	
(e)	HM			*	*	*

c. **Tableau for HM**

	HM	END[S]	DIFF[T]	END[T]	SAME[S]	SAME[T]
(a)	M	*!			*	*
(b)	ML	*!	*	*	*	
(c)	MLh	*!		**	*	*
(d)	Lh					*
(e)	HM		*!	*		

d. **Tableau for Lh**

	Lh	END[S]	DIFF[T]	END[T]	SAME[S]	SAME[T]
(a)	M	*!	*		*	
(b)	ML	*!		*	*	*
(c)	MLh	*!		**	*	*
(d)	Lh		*!			
(e)	HM			*		*

The remainder of the Eastern A-Hmao tone system is very much like that of Western A-Hmao, and can be modelled in much the same manner.

The A-Hmao circle-shift cases are of great value in that they demonstrate that a phenomenon which has sometimes been viewed as a parochial property of Southern Min languages has arisen independently in a language family that is not in contact with Southern Min dialects. Of course, Min dialects and Far-Western Hmongic languages have a few things in common: both, for example, have relatively large tonal inventories

and both groups have relatively extensive tone sandhi including sandhi of the structure building (non-structure preserving) type. The crowded tone space that results supporting so many contrasts facilitates mergers of allotones and sandhi tones with tones in the inventory other than their underlying tone, resulting in chain shifts and, under the right conditions, circle shifts. However, as the next case will show, tonal circle shifts can arise without a large tonal inventory.

4.3 Jingpho Bounce-back Effect

Jingpho presents one of the most intricate and fascinating cases of circular chain shifting (specifically, bounce-back) currently known. The issue of circularity in Jingpho's tonal grammar was first raised by Lai (2002) and was subsequently discussed and re-analyzed by Mortensen (2003). The analysis presented here, while drawing on aspects of those earlier analyses, is applied to a larger set of data which better represent the complexity of Jingpho tone sandhi.

The alternations discussed here, based on data from Dai (1990b) and Lai (2002), occur when two (monosyllabic) nominal roots are concatenated to form a subordinating (attributive) compound. As will be seen, all of these alternations driven by a requirement that the surface correspondent of the underlying initial tone be different from it—that is, at a different point on the scale. This results, as we will see, in situations where H becomes L in the same environment where L becomes H, or L becomes M in the same environment where M becomes L.

An important fact about the tonal patterns in Jingpho compounds is that the sandhi patterns are totally different in cases where the first syllable is CHECKED (ends in a stop coda) and cases where the first syllable is SMOOTH (is open or ends in a sonorant coda). Because it is simpler, we will begin with a discussion of the patterns associated with initial checked syllables, after which we will provide an analysis of compounds

with initial smooth syllables—a more complicated pattern.

4.3.1 With initial checked syllables

In smooth syllables, there is a three-way tonal contrast in Jingpho between H (high), M (mid), and L (low). However, in checked syllables, there is only a two-way contrast between L and H. The outputs corresponding to all the possible combinations of input tones is given in (70):

(70) **Tone sandhi in compounds with initial checked syllables**

$\sigma_1 \backslash \sigma_2$	H	M	L
H	LL	LL	LL
L	(HH)	HM	HL

The generalization seems to be as follows: the initial tone must be different in the input and the output. Since there are only two options, this means that L becomes H and H becomes L. Furthermore, the contour of the tones of the two syllables taken together must not have a rising slope, so the second tone in the H+H and H+L combinations must be L in the output so that they will be no higher than the L in the first syllable. Tones are changed minimally in order to achieve these ends.

Examples illustrating these patterns are given below:

(71) **H+H → LL**

- a. /júp tám/ [júp tám]
‘sleepwalking’
- b. /wáʔ sáj/ [wàʔ sàŋ]
‘XX’

(72) **H+M → LL**

- a. /jít khɜai/ [jìt khɜài]
‘patch of dry land’
- b. /ŋáʔ li/ [ŋàʔ lì]
‘banana leaf bud’

(73) **H+L** → **LL**

a. /pʒáʔ mùt/ [pʒàʔ mùt]

‘gray caterpillar’

b. /lút jòm/ [lùt jòm]

‘cigarette’

(74) **L+M** → **HM**

a. /mjìʔ pʒuì/ [mjíʔ pʒuì]

‘tear’

(75) **L+L** → **HL**

a. /nàʔ thàm/ [náʔ thàm]

‘kind of large taro’

b. /wàʔ kʒàŋ/ [wáʔ kʒàŋ]

‘bamboo wash basin’

If we assume the scale T as in (4.3.1), we can capture our generalization quite simply in terms of the constraints that were proposed in Chapter 2.

(76) $T = \{L\} < \{M\} < \{H\}$

The imperative that the initial tone change can be expressed in terms of the constraint $\text{DIFF}[T]_{\# \sigma}$, which states that the tone in the output corresponding to the initial tone in the input must be different (in T) from its correspondent in the input. This constraint must outrank another constraint, $\text{NOWAX}[T]$, which expresses our generalization that the second tone may not be higher than the first. In order to know what effect NOWAX will have, we need to talk about the ranking of $\text{CORR-}\sigma - \sigma$, the constraint that enforces correspondence between adjacent syllables (to a first approximation)⁴. What we must say about it is that it is ranked above the other constraints we have employed, meaning that the syllables—and therefore tones—of adjacent syllables are always in correspondence.

Finally, we must assume that there is a high-ranking constraint that bans non-peripheral tones from occurring in checked syllables, namely $\text{EXTREME}[T]$, and a non-

⁴There should also be another constraint on correspondence relationships that penalizes relationships between syllables that are not next to each other. However, since the cases we are dealing with here only have two syllables, that fact is not crucial to this analysis.

mal scalar identity constraint, SAME[T]. EXTREME[T] should be undominated and SAME[T] should be ranked below the other three constraints. This ranking is capable of generating all of the tonal patterns seen in Jingpho checked syllables:

(77) a. **Jingpho checked H plus L**

/H+L/	EXT[T] _q	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]
(a) HL			*!	
☞ (b) LL				*
(c) LH		*!		**
(d) ML	*!			*

b. **Jingpho checked L plus L**

/L+L/	EXT[T] _q	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]
(a) LL			*!	
(b) ML	*!			*
☞ (c) HL				*
(d) HM				**!

c. **Jingpho checked H plus H**

/H+H/	EXT[T] _q	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]
(a) HH			*!	
☞ (b) LL				**
(c) LH		*!		*
(d) ML	*!			**

d. **Jingpho checked H plus M**

/H+M/	EXT[T] _q	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]
(a) HM			*!	
(b) LL				**
(c) LM		*!		
(d) MM	*!			

The four tableaux in (77) show how the “neutralization with bounce-back” pattern that we can see in Jingpho checked syllables is generated by the grammar. Before L, L becomes H but H becomes L, as we have noted previously. Thus, in this environment there is not neutralization. However, this pair of alternations participates in a larger pattern of neutralization if the tones of both syllables take taken together. Looking at the data from this perspective, it is clear that H+M and H+L are neutralized with H+H as [LL], while L+L becomes [LH]. The grammar we have developed here helps us to see why this is the case. The neutralization to [LL] occurs because it is the only way to avoid ascending the tone scale across a word if the initial tone is L. Since underlying H in initial position *must* become [L] in the output, then any sequence of H plus some tone will be realized as [LL]. Underlying LL, however, cannot be realized as [LL] since this would mean no change in the scale position of the first tone. It cannot become M, because M tones are banned in checked syllables, so its only option is to bounce-back to [HL].

4.3.2 With initial smooth syllables

Peculiarly, in compounds where the first syllable is smooth, that is, does not have a stop coda, the tonal patterns are strongly antagonistic to those in compounds with checked syllable initially. In fact, they behave as if the tone scale has been inverted, with L at the top and H at the bottom. If we make this change, and leave the other constraints

the same, we come very close to modelling the pattern in smooth syllables. We will leave aside the question of why this should be the case and concentrate on modelling the alternations, demonstrating the how scales and anti-identity constraints are able to capture the rather intricate pattern we see in Jingpho tone sandhi in an elegant and insightful fashion.

The basic patterns seen in Jingpho tone sandhi (with smooth syllables in the initial position) is given in (78):

(78) **Jingpho tone patterns: disyllables with smooth σ_1**

$\sigma_1 \backslash \sigma_2$	H	M	L
H	LH	LM	LL
M	LH	LM	LL
L	HH	MM	LH

Data illustrating these patterns are given below:

(79) **H+H → LH**

- a. /túm khʒák/ [tùm khʒák]
‘small box’
- b. /phún kúʔ/ [phùn kúʔ]
‘crooked tree branch’

(82) **M+H → LH**

- a. /khon ʒám/ [khòn ʒám]
‘young girl’
- b. /khon sét/ [khòn sét]
‘clever girl’

(80) **H+M → LM**

- a. /túm kə/ [tùm kə]
‘variegated tube’

(83) **M+M → LM**

- a. /khʒaŋ kho/ [khʒàŋ kho]
‘cabbage’

(81) **H+L → LM**

- a. /nói tʃi/ [nòi tʃi]
‘small basket’

- b. /si phʒa/ [sì phʒa]
‘cotton field’

- (84) **M+L → LL**
- a. /lɔŋ phɔ̀/ [lɔŋ phɔ̀]
‘white clothing’
- b. /juŋ fà/ [juŋ fà]
‘little finger’
- (85) **L+H → LM**
- a. /ù ɔ́/ [ù ɔ́]
‘pigeon cage’
- b. /tɪŋ tsóŋ/ [tɪŋ tsoŋ]
‘XX’
- (86) **L+H → HH**
- a. /thìŋ ɔ́ʔ/ [thíŋ ɔ́ʔ]
‘small bamboo tube’
- b. /nà khám/ [ná khám]
‘wasteland’
- (87) **L+M → MM**
- a. /thìŋ ka/ [thɪŋ ka]
‘back basket’
- b. /ù khɔ́/ [u khɔ́]
‘turtle dove’
- (88) **L+L → LH**
- a. /kìn thòŋ/ [kìn thóŋ]
‘XX’
- b. /sìn tɪn/ [sìn tín]
‘diaphragm’

The basic pattern is easily summarized: H and M are neutralized to L in all environments. L assimilates to the following tone before H and M. The complication comes in compounds with an underlying LL sequence, which surface with a LH melody. This seems exceptional in two respects. First, the tone of the first syllable does not seem to change in this case, though it changes in every other case. We will see shortly that this is actually something of an illusion—a result of tonal metathesis. It is important to point out, before proceeding to a technical analysis, that the tonal patterns contain two bounce-back patterns. First HH and MH become LH but LH becomes HH; second, HM and LM become LM but LM becomes MM. This forms some, but certainly not the only, evidence for the anti-identity constraints at work in this phenomenon.

There is another crucial generalization that must be made in order to get the analy-

sis of Jingpho tone sandhi right which is identical (modulo the sequence of the scale) to an observation already made with the checked syllables: there are no tonal melodies with a falling contour, pitch-wise (and thus, a rising contour, scale-wise). This indicates that NOWAX[T] is active in this system. Since it is (almost) always the case that it is the tone of the first syllable which changes, we must also assume that DIFF[T]_{#σ} is active. Following the argument from the checked syllables, we will assume that it is dominated by NOWAX[T] and that it dominates SAME[T]. There must also be some explanation for the fact that underlying H and M are neutralized to L in all environments, a statement—in effect—that L is at the “unmarked” end of the scale. Since L is at the top of the scale, we must conclude that TOPMOST[L] dominates ENDMOST[L]. All of these ranking statements are either identical or perfectly consistent to those motivated for the checked-syllable compounds.

Given this ranking alone, we are able to generate the correct outputs for inputs with initial H:

(89) **Jingpho smooth H plus H**

	/H+H/	NOWAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	HH		*!		**
(b)	LH			*	*
(c)	MH			*	**!
(d)	LL			*!*	

(90) **Jingpho smooth H plus M**

	/H+M/	NOWAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	HM		*!		**
(b)	MM			*	**!
(c)	LM			*	*
(d)	LL			*!*	

(91) **Jingpho smooth H plus L**

	/H+L/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	HL		*!		*
(b)	ML			*	*!
(c)	MM			*!*	**
☞ (d)	LL			*	

DIFF[T]_{#σ} compels a change in the tone of the initial syllable, provided this can be done while still satisfying NOwAX[T], and this is trivially the case. This change will always be to L since that will minimize violations of TOPMOST[L] without incurring violations of other constraints. The same principles hold for outputs with initial M:

(92) **Jingpho smooth M plus H**

	/M+H/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	MH		*!		**
(b)	HH			*	**!
☞ (c)	LH			*	*
(d)	LL			*!*	

(93) **Jingpho smooth M plus M**

	/M+M/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	MM		*!		**
(b)	HM	*!		*	**
☞ (c)	LM			*	*

(94) **Jingpho smooth M plus L**

	/M+L/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	ML		*!		*
(b)	HL			*	*!
(c)	LL			*	
(d)	LH	*!		**	*

However, for outputs with initial L, this set of constraints is not adequate, as can be seen in the following tableaux. For /L+H/ and /L+M/, the grammar badly under-determines the output because it has no access to the generalization that, other things being equal, melodies in which the two tones are identical are preferred:

(95) **Jingpho smooth L plus H**

	/L+H/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	LH		*!		*
(b)	LL		*!	*	
(c)	MH			*	**
(d)	HH			*	**
(e)	MM			**!	**

(96) **Jingpho smooth L plus M**

	/L+M/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	LM		*!		*
(b)	HM			*	**
(c)	MM			*	**
(d)	HH			**!	**

The grammar is even worse for the input /L+L/, for which the output is not only badly underdetermined, but incorrectly determined as well. This is because the only repairs that allow the satisfaction of both NOwAX[T] and DIFF[T]_{#σ} (at least the only repairs we are considering at this point) involve raising both tones. The first tone is raised

since that is the only way to satisfy $\text{DIFF}[T]_{\#\sigma}$; the second is raised because in order to satisfy $\text{NOWAX}[T]$ it must be as high or higher in pitch (that is, as low or lower on the scale) than the first.

(97) **Jingpho smooth L plus L**

	/L+L/	NOwAx[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]
(a)	LL		*!		
(b)	ML	*!		*	*
(c)	HL	*!		*	*
☹	(d) MH			**	**
(e)	LH		*!	*	*
(f)	LM		*!	*	*
☹	(g) HH			**	**
☹	(h) MM			**	**

For /L+H/ and /L+M/, the problem can be fixed by adding the constraint $\text{PLATEAU}[T]$ which distinguishes between the candidate MH and the “level” outputs HH and MM. This constraint does not need to be high-ranked—only to exist somewhere in the constraint hierarchy, as demonstrated by the following tableaux:

(98) **Jingpho smooth L plus H**

	/L+H/	NOwAx[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]
(a)	LH		*!		*	*
(b)	LL		*!	*		
(c)	MH			*	**	*!
☹	(d) HH			*	**	
(e)	MM			*!*	**	

(99) **Jingpho smooth L plus M**

	/L+M/	NOwAx[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]
(a)	LM		*!			*
(b)	HM			*	**	*!
(c)	MM			*	**	
(d)	HH			*!*	**	

It is worthy of note that PLATEAU[T] must be dominated by TOP[T]. Otherwise, inputs like /H+M/ and /M+H/ would be mapped to MM and HH, respectively. PLATEAU[T] only comes into effect when it is impossible to satisfy DIFF[T]_{#σ} by outputting a tone at the top of the scale (that is, L). Ranking PLATEAU[T] here encodes the insight that the next-best thing to being L for a tone is to agree with the other tone in the compound.

This ranking is correct for /L+H/ and /L+M/, but it is not adequate for /L+L/. On the one hand, it narrows down the number of tying candidates by one; on the other hand, neither of these candidates is the correct output:

(100) **Jingpho smooth L plus L**

	/L+L/	NOwAx[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]
(a)	LL		*!			
(b)	ML	*!		*	*	*
(c)	HL	*!		*	*	*
(d)	MH			**	**	*!
(e)	LH		*!	*	*	*
(f)	LM		*!	*	*	*
⊖	(g) HH			**	**	
⊖	(h) MM			**	**	

Under this ranking HH and MM beat MH because, while they otherwise incur the same violations, HH and MM satisfy PLATEAU[T] while MH does not. This is not really progress, though, since the desired output is LH. This looks difficult—even impossible—to generate under our current ranking. The problem, though, is not the

ranking as much as it is the assumptions we have been making about the tones. If we treat them as autosegments, there is no reason—within OT—to believe that the linear ordering of tones in the output might not be the same as the ordering of tones in the input. We ought to suppose, as others have done in the past, that there is a family of faithfulness constraints, LINEARITY (abbreviated LINEAR), that penalize candidates in which the linear order of entities is not the same as that in the input. If we assume that linearity is relatively low-ranked, “metathesis” or changes in the linear order of autosegments, is available as a repair strategy for satisfying higher-ranked constraints.

Applying these new assumptions to /L+L/, we get the following result:

(101) **Jingpho smooth L plus L, allowing for non-linearity**

/L _i +L _j /	NOwax[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR
(a) L _i L _j		*!				
(b) M _i L _j	*!		*	*	*	
(c) M _j L _i	*!		*	*	*	*
(d) H _i L _j	*!		*	*	*	
(e) M _i H _j			**!	**	*	
(f) L _i M _j		*!	*	*	*	
⌘ (g) L _j M _i			*	*	*	*
(h) L _i H _j		*!	*	*	*	
⌘ (i) L _j H _i			*	*	*	*
(j) H _i H _j			**!	**		
(k) M _i M _j			**!	**		

The grammar now selects two outputs, LM and LH—one of which is correct. This is a considerable step forward. What it indicates is that the grammar has another characteristic: other things being equal, M is dispreferred. This can be expressed via the constraint EXTREME[T], which penalizes tones that are not at the end-points of the scale. Once EXTREME[T] (abbreviated EXT[T]) is added to the ranking, the correct output is selected:

(102) **Jingpho smooth L plus L, adding EXTREME[T]**

	/L _i +L _j /	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	L _i L _j		*!					
(b)	M _i L _j	*!		*	*	*		*
(c)	M _j L _i	*!		*	*	*	*	*
(d)	H _i L _j	*!		*	*	*		
(e)	M _i H _j			**!	**	*		*
(f)	L _i M _j		*!	*	*	*		*
(g)	L _j M _i			*	*	*	*	*!
(h)	L _i H _j		*!	*	*	*		
(i)	L _j H _i			*	*	*	*	
(j)	H _i H _j			*!*	*			
(k)	M _i M _j			*!*	**			**

It is comforting to note that this same ranking makes the correct predictions for the cases we examined earlier. For the cases with initial H or M, linearity is not an issue, since it is always possible to satisfy both NOwAX[T] and DIFF[T]_{#σ} simply by changing the first tone to L. For these cases, then, we will not look at linearity, but will provide tableaux with the same candidates as before, demonstrating that the constraint ranking is still valid:

(103) **Jingpho smooth H plus H**

	/H+H/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	HH		*!		**			
(b)	LH			*	*	*		
(c)	MH			*	**!	*		*
(d)	LL			*!*				

(104) **Jingpho smooth H plus M**

	/H+M/	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	HM		*!		**	*		*
(b)	MM			*	**!			**
(c)	LM			*	*	*		*
(d)	LL			*!*				

(105) **Jingpho smooth H plus L**

	/H+L/	NoWax[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	HL		*!		*	*		
(b)	ML			*	*!	*		*
(c)	MM			*!*	**			**
(d)	LL			*				

(106) **Jingpho smooth M plus H**

	/M+H/	NoWax[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	MH		*!		**	*		*
(b)	HH			*	**!			
(c)	LH			*	*			
(d)	LL			*!*				

(107) **Jingpho smooth M plus M**

	/M+M/	NoWax[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	MM		*!		**			**
(b)	HM	*!		*	**	*		**
(c)	LM			*	*	*		*

(108) **Jingpho smooth M plus L**

	/M+L/	NoWax[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a)	ML		*!		*	*		*
(b)	HL			*	*!	*		
(c)	LL			*				
(d)	LH	*!		**	*	*		

The situation is slightly more complicated for /L+H/ and /L+M/ in that linearity is not trivially unimportant in these cases. However, as shown in the following tableaux, the ranking as we now have it will generate the correct outputs even when linearity is taken into account:

(109) **Jingpho smooth L plus H, considering linearity**

$/L_i+H_j/$	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a) L_iH_j		*!		*	*		
(b) L_iL_j		*!	*				
(c) M_iH_j			*	**	*!		*
(d) H_iH_j			*	**			
(e) H_jH_i			*!	**		*	
(f) M_iM_j			*!*	**			

(110) **Jingpho smooth L plus M, considering linearity**

$/L_i+M_j/$	NOwAX[T]	DIFF[T] _{#σ}	SAME[T]	TOP[T]	PLAT[T]	LINEAR	EXT[T]
(a) L_iM_j		*!			*		*
(b) H_iM_j			*	**	*!		*
(c) M_iM_j			*	**			**
(d) M_jM_i			*!	**		*	**
(e) H_iH_j			*!*	**			

We have demonstrated, then, that the grammar at which we have arrived can generate all of the input-output tone mappings for Jingpho noun compounds having an initial smooth syllable. However, in arriving at a grammar that properly characterizes the smooth-syllable compounds, we have changed it so that it is no longer completely adequate for the checked-syllable compounds. Specifically, we have posited that LINEARITY is dominated by NOWAX[T]. In doing this, we allow for the possibility that sequences with an initial H underlyingly could be repaired in the output (where that H must become L) so that they satisfied NOWAX[T] by metathesis of the tones, rather than by lowering. This is as puzzling an aspect of the tonal grammar of Jingpho as is the “scale-reversal” that holds between the smooth and checked compounds. However, it can be solved if we assume that there is a high ranked LINEARITY constraint that becomes active just in case the first syllable of the compound is checked. This could be accomplished through local constraint conjunction. Another solution which would not require conjunction would be a FAITH-ASSOC constraint that applied specifically to checked syllables, requiring that tones, as entities rather than properties, be associ-

ated with the same syllable in input and output if that syllable has a stop coda. The problem with this approach is that it predicts a different behavior for compounds with checked syllables in the second position. This does not seem to accord with the data. This question cannot be fully resolved here, and is a avenue for further research.

What has been established is that Jingpho tone sandhi includes “bounce-back” patterns; that is, minimal circle shifts and that these patterns and the larger set of tone patterns in Jingpho can be modelled insightfully with the formalism of scales and scale-referring constraints that were proposed in Chapter 2 and developed further in Chapter 3. This means that there are a number of different cases of circular chain shifts in tone languages of Asia, and that some of these are necessarily independent of the others, given the geographical barriers that separate the languages in which they appear. This means that circular chain shifts are more significant (theoretically speaking) and less bizarre (empirically speaking) than phonologists have previously believed.

4.4 Circle Shifts and Bounce-back effects as counter-examples to harmonic ascent

Circle shifts touch on a crucial issue in Optimality Theory and phonological theory in general, namely the property that has been dubbed by Moreton (2004b) *Harmonic Ascent*:

(111) Harmonic Ascent

An unfaithful mapping from input to output only occurs to satisfy a markedness constraint which dominates the faithfulness constraints that would militate against that mapping. *All unfaithful mappings decrease markedness.*

Through a corrected and rigorous formal proof, Moreton showed that what he called “classical Optimality Theory” had this property. By classical Optimality Theory, he meant a theory of grammar in which there were only markedness and faithfulness

constraints: markedness constraints penalizing certain structures, and faithfulness constraints that enforced similarity (and only similarity) between inputs and outputs. What Moreton demonstrated, and called “Harmonic Ascent,” was that such grammars can only generate certain types of mappings—any type of mapping, in fact, other than circular chain shifts and infinite chain shifts. The logical reasons for this are quite simple: the only way for unfaithful mappings (that is, mappings where the input differs from the output) to take place in such a model is for the output to be less marked than the input, relative to the hierarchy of markedness constraints in the grammar. If some input A that is mapped to a non-identical output B, then it logically follows that B is less marked than A. Thus, there is no reason, in classical Optimality Theory, for an input like B ever to be mapped to an output like A. A similar, simple logic holds for infinite chain shifts, which will not be of immediate concern here.

Harmonic Ascent stands as one of the great formally-validated predictions of Optimality Theory, and there has been great reluctance to introduce theoretical devices that would undermine it, given that it is not clear that there are other, similarly far-reaching predictions that follow from the OT formalism. Łubowicz (2003) was careful to design her PC Theory so that it had this property (or appeared to have this property). Barrie (2006), however, seems to have shown that at least one form of PC Theory is not, in fact, subject to Harmonic Ascent. McCarthy (2002), likewise, presents as one argument against his own theory of comparative markedness the fact that it does not obey Harmonic Ascent and can actually generate both infinite chain shifts and at least one type of circle shift (though it does not seem able to generate circle shifts with more than two steps, like those found in some Southern Min languages).

Moreton was, of course, aware of the obvious empirical objections to a theory of phonology that obeyed Harmonic Ascent, specifically the existence of the Southern Min tone circle. In discussing this phenomenon, and why it should not be taken as a

true counter-example to Harmonic Ascent, Moreton presents two claims: The first is that the Xiamen/Mainstream Taiwanese tone circle was not the kind of phonology a theory like OT was meant to model, but is, rather, what Schuh (1978) called it, namely “paradigmatic replacement.” In other words, the process simply represents the arbitrary replacement of one tone in the inventory with another. His other claim that he presented was that the tone circle was actually morphological, or represented an interface process between the phonology and the morphology and thus would not be subject to Harmonic Ascent.

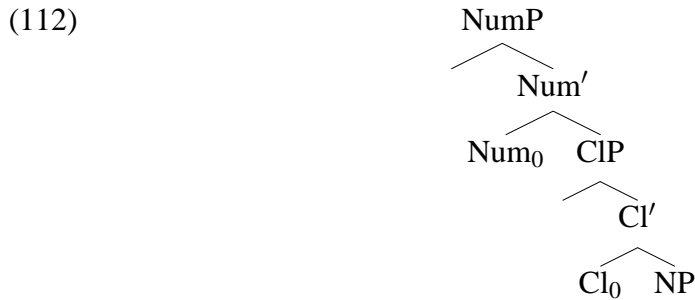
Moreton’s first claim is not entirely in disharmony with the claims made in this dissertation: the system of scales and constraints that are introduced here can produce a kind of “paradigmatic replacement.” I innovate here, however, by showing that these replacement processes are not an isolated phenomenon in the sound patterns of languages, but rather are linked to a whole host of phenomenon, some of which cannot be readily exiled from the realm of “normal phonology.” The second claim presents more problems. If the claim is that the tone circle is not subject to Harmonic Ascent because it involves the interface between phonology and morphology (and it is by no means clear that it does depend on such a relationship in any special way), then we should expect any process that involve an interaction between phonology and morphology to be potentially independent of Harmonic Ascent. Since this is true, by definition, of all morphophonology, Harmonic Ascent should only be a property of low-level phonological processes of the allophonic type. If this is held to be true, then there is no principled reason that morphophonology should necessarily be subject to *any* predictions made by the theory of phonology. This position has its adherents, to be sure, and is perfectly valid on its own terms. However, it is not consistent with the goals that theoretical phonologists have been trying to reach since the advent of generative phonology.

This type of philosophical observation becomes especially important when we look more closely at the relationship between morphosyntax and phonology in languages displaying circle shifts. Certainly, in Jingpho, the alternations are closely tied to a particular morphological context, namely noun-noun compounds. In the Southern Min cases, it has been held that the sandhi takes place when a syllable is non-final in some XP, though this claim is by no means uncontroversial, and some investigators hold that the conditioning environment is, in fact, phonological (Chen 1987). We have also looked at the A-Hmao cases without discussing the relationship between these phonological processes and morphosyntax. It is my claim that the domain for these processes is not, in fact, morphosyntactic (as it may first appear) but is actually prosodic, sandhi contexts being confined to prosodic words.

Tone sandhi in A-Hmao (Eastern and Western) behaves very much like that in other Far Western Hmongic languages. Tone sandhi is found in three principle environments: within verb phrases, specifically within verb-noun compounds; in noun phrases, particularly in noun-noun compounds or phrases where a noun is modified by a stative verb; and (crucially, for our argument) between numerals and classifiers, with numerals being the triggers and classifiers being the targets (Downer 1967; Heimbach 1979; Wang and Wang 1986; Ratliff 1992b; Niederer 1998). That verb-noun compounds, noun-noun compounds, or nouns with verbal modifiers (possibly best treated as compounds as well) are morphosyntactic constituents is uncontroversial. Were these the only places where sandhi alternations were present, it would be ambiguous whether the domain for these processes was prosodic or morphosyntactic (since they are both morphosyntactic constituents and form prosodic words). It is less clear, however, that numerals and classifiers together form morphosyntactic constituents.

There is a position emerging among syntacticians who study Hmongic languages and other languages with similar syntactic structures, that both classifiers and numerals

head their respective phrases within the DP, with the NP nested inside the classifier phrase (CIP) and the classifier phrase nested inside the numeral phrase (NumP) (Simpson 2005; Sio 2006). This may be represented graphically, leaving out all projections other than those with which we are immediately concerned, as follows:



Classifiers typically do not trigger tone sandhi in nouns that they precede⁵. However, the Num cannot form any constituent with the Clf to which the NP does not belong. What does link the Num and Clf together is the fact that they are parsed together in phrasing, as a single prosodic word. There is a mismatch, then, between prosodic structure and morphosyntactic structure, and tone sandhi ignores the morphosyntax in deference to prosody. This is important in understanding the true significance of the tone sandhi circles in A-Hmao: circle shifts are not dependent on morphosyntactic conditioning and thus cannot be denied the same status as other phonological processes on principled grounds, even if morphologically-conditioned phonology is off-limits to proper phonological theory.

This chapter should help address another objection that has been advanced by Moreton (2004b) and others, namely, that circle shifts are confined to a single family and are thus *sui generis*—a curiosity, but not the business of phonological theory. As I have

⁵The exceptions to this seem to be in DPs where there is no overt numeral. It appears that the classifier is phonologically dependent and will be grouped with the numeral where one is present but with the noun when there is no numeral. This supports the idea that Hmong classifiers behave phonologically like clitics.

shown here, there are multiple different tone circles in Southern Min, some of which must have developed independently of the others, indicating that the circle was not just some historical hold-over but that new circles can emerge of their own accord under the right conditions. Probably more important, this chapter has provided detailed analyses of two (or, if Eastern and Western A-Hmao are counted separately, three) cases of circle shifts that are not found in Southern Min dialects. This should dispel the idea that circle shifts are limited to a single small linguistic subgroup and thus need not be seen as a general linguistic phenomenon.

The ultimate conclusion of this discussion must be that Moreton's Theorem, while it is undoubtedly true, is true of a theory that is descriptively and explanatorily inadequate. We have seen that Structure Optimality and the theory of logical scales are superior to both classical Optimality Theory and contrast preservation theories in describing and accounting for circle shifts. As has already been seen in Chapter 3, and as will become more clear in the following chapters, the same formal tools that were used in analyzing circle shifts are required for accounting for other phenomena. As a result, we are able to see circle shifts not as bizarre patterns far-removed from other phonological phenomena, but as a natural consequence of a general theory of phonological grammar. In the next chapter, for example, we will examine ordering effects in coordinate compounds, a phenomenon that can best be explained in terms of the same kind of directional anti-identity that was invoked in this chapter to model circle shifts.

Chapter 5

Ordering Effects in Coordinate Compounds

Primus inter pares

Roman imperial title

Some of the most interesting evidence regarding the existence and nature of phonological scales comes from phonological constraints on the ordering of coordinate compounds. While the existence of such effects has long been known, their significance has largely been missed by phonologists, possibly because of the lack of a framework by which these constraints could be correlated with the better-understood patterns in the phonologies of the world's languages. The scale formalism that has been developed in Chapter 2 and applied to chain-shifts and circle shifts in Chapters 3 and 4 will be employed here as such a framework. On the one hand, it will be shown that logical scales are able to account for otherwise difficult phenomena in coordinate compound ordering; on the other hand, it will be seen that these phenomena reinforce and refine what has been said thus far regarding the properties of these scales.

This chapter begins, in §5.1, with a survey of coordinate compounding and related phenomena, which provides an empirical context to the analytic discussion that follows. This is followed, in §5.2, by an overview of the types of ordering constraints that may affect coordinate compounds. Taking this background information as a starting point, the following sections comprise specific case studies and analyses of ordering effects from a number of different languages: Chinese in §5.3, Lahu in §5.4, Jingpho and Tangkhul in §5.5, Hmong in §5.6, Qe-Nao in §5.7 and Qo-Xiong in §5.8.

5.1 Coordinate Compounds

COORDINATE COMPOUNDS (henceforth *co-compounds*) are a class of compound words which are “headless” but not exocentric, that is to say none (or more typically, neither) of their constituents act individually as the head of the compound, but they also are not simply modifiers for an external head. In the broad sense, English copulative (dvandva) compounds like *poet-prophet* or *singer-songwriter* are coordinate compounds. A *poet-prophet* is not a prophet of the poetic ilk, nor, indeed, is he a poet of the prophet stripe. *Poet-prophets* are exactly that set of beings who are both poets and prophets. For this reason, the sequence of the elements can be changed to *prophet-poet* without changing the meaning of the expression. Such compounds are clearly different from exocentric compounds like *pickpocket* ‘person who picks pockets’, *kill-joy* ‘person who kills joy’, or *spoilsport* ‘person who spoils sport’ where the referent that is being modified or specified is not named by either of the constituents of the compound. While neither of the constituent roots in an exocentric compound name the referent, *both* of the consistent roots in a coordinate compound name the referent.

In copulative compounds, the effect coordination is to further restrict what the compound can refer to. The set of entities that are *poet-prophets* is a subset of both the set of entities that are poets and the set of entities that are prophets. This is not the only possibility, however. In many languages, there are coordinate compounding constructions with generalizing semantics. In Mong Leng, the historical root meaning ‘face’ has been completely replaced by a compound of ntjê ‘ear’ and mja ‘eye’. The meaning of the resulting compound, ntjê-mja, is not ‘things that are both ears and eyes’, nor simply ‘the set of things that are ears or eyes’, but ‘ears, eyes, and so forth.’ Because ‘ear’ and ‘eye’ are representative parts of the face, a compound of these two words can refer to the parts of the face as a class (and, for practical purposes, the whole face). Such compounds may also be composed of synonymous or nearly-synonymous parts.

For example, in Jingpho *məsùʔ* ‘deceive’ and *khólém* ‘fool (v.)’ may be compounded to form *məsùʔ-khólém*. In these cases, no additional principle of interpretation need be invoked; if a generalizing coordinate compound refers to the smallest natural class including the referents of both constituents, the meaning of a compounded pair of synonyms is predicted to be approximately the same as that of each of the parts, modulo whatever pragmatic information is conveyed by a coordinate-compounding construction in the language in question.

Compounds of this type can be compared to two different types of English constructions. The first are the very small set of echo-reduplication constructions, such as *flip-flop*, *tip-top*, *drip-drop*, and *sing-song* (Marchand 1969). The structure of such words—and, indeed, what roots can enter into such constructs—are constrained by a phonological template (an idea the importance of which will become evident in Section 5.5 below). The other English construction that has characteristics in common with generalizing coordinate compounds are what Malkiel (1959) called “irreversible binomials.” These are lexicalized constructions, usually including a coordinating conjunction, such as *kith and kin*, *ways and means*, or *warp and woof*. These constructions may have generalizing semantics as well. For example, *trials and tribulations* seems to refer not just to tests and to causes of suffering, but to difficult and unpleasant experiences generally. Likewise, *odds and ends* are not simply things that are without mates or too short to use, but to assorted (prototypically left-over) items as a class. As will be discussed below, there seem to be both semantic and phonological tendencies which influence the sequence of elements in idioms of this type (Cooper and Ross 1975). However, recent investigators like Benor and Levy (2006) have argued that phonological constraints are far less important than semantic constraints in determining the order of conjuncts in binomials of this kind.

5.1.1 Co-compounds in Mainland Southeast Asia

Most languages of mainland Southeast Asia have a class of binary coordinate compounds, typically of the generalizing or additive type. That, by itself, is not especially significant—many languages feature constructions of this sort. What is striking about languages of mainland Southeast Asia is the important role that coordinate compounds play in their lexicons. A good example is Hmong: Co-compounds occur with high frequency in all speech registers and the primary (or only) words for many basic concepts, like ‘food’ ʒáu-mǔ, ‘face’ ntjê-mũa, ‘land’ té-chài, ‘time’ câi-ᵛᵛᵛ, or ‘clothing’ t̩-t̩hò, are coordinate compounds. Many of these words are common enough that they are learned as units, and some speakers do not seem aware that they can be analyzed into meaningful parts until this is brought to their attention. However, most compounds of this type are analyzable and the construction, as a whole, is quite productive. New compounds can be formed using both native and borrowed vocabulary. As Hmong in peninsular Southeast Asia started entering the modern economy, various words for wage work were coined, among them lâ-kàj ‘job (“business-work”)’. The second element of this compound is borrowed from a Daic language and is cognate to Thai ᵛᵛᵛ kaan ‘work; job; task; business’. This compound is apparently a recent coinage. A large number of novel compounds, though, are basically ephemeral and do not come to reside in the lexicon.

In the lexicons of many other languages of East and Southeast Asia, co-compounds play just as important a role as they do in Hmong. In Vietnamese, for example, many rather fundamental concepts are expressed through compounds of this kind (Thompson 1965:128):

- (113) a. quần - áo
 pants tunic
 ‘clothing’
- b. bàn - ghế
 table chair
 ‘furniture’

c. bát - đĩa
bowl plate
'dinnerware'

d. mua - bán
buy sell
'go shopping'

In this, Vietnamese is like other Mon-Khmer languages. In Khmer, there are a large quantity of co-compounds, many of them composed of rhyming or alliterative pairs (Ourn and Haiman 2000:486):

(114) a. baac - saac
throw sprinkle
'shed, scatter, throw'

c. bat baj
disappear lose
'disappear'

b. baok - baen
beat trample
'thresh'

d. priəp thiət
compare compare
'metaphorical'

Pacoh, also a Mon-Khmer language, employs co-compounds quite freely, allowing "modifiers" and verbs, as well as nouns, to participate in this type of construction (Watson 1966):

(115) **Nouns**
a. qachat - qakốq
axe bushknife
'tools'

b. kểaq - chư
guard watch
'take care of'

b. qaqât - qachếq
animals birds
'wildlife'

(117) **Modifiers**

a. hây - tubếq
able wise
'capable'

(116) **Verbs**

a. klổn - qanhờy
play play
'play'

b. qian - qo
easy good
'peaceful'

In all of these cases, co-compounds are far from marginal. To the contrary, they are among the most important word-formation strategies in each of these languages.

Likewise, in Chinese, co-compounds are an exceptionally prominent part of the lexicon, ranging in their semantics from completely transparent to highly opaque:

- | | | |
|-------|--|--|
| (118) | a. 天 <i>tiān</i> - 地 <i>dì</i>
heaven earth
'universe' | c. 死 <i>sǐ</i> - 活 <i>huó</i>
die live
'fate' |
| | b. 長 <i>cháng</i> - 短 <i>duǎn</i>
long short
'length' | d. 風 <i>fēng</i> - 水 <i>shuǐ</i>
wind water
'geomancy' |

While the importance and frequency of compounds of this type seem to reach their apogee in the languages of China and mainland Southeast Asia, they are found in many other languages and language families. Some particularly comparable examples can be found in Sumerian, a language isolate of ancient Mesopotamia. In Sumerian 𒀭 𒂗 AN.KI *an-ki* 'heaven-earth' means 'universe', in a matter perfectly comparable to Chinese 天地 *tiān-dì* 'universe ("heaven-earth") or Hmong ntû-té 'id.' Likewise, Sumerian 𒀭 𒂗 MAŠ.ANŠE *maš-anše* 'goat-donkey' means 'animals' or 'livestock' (Tinney 2006)¹. Both the binary structure and the generalizing semantics of these compounds seem to be identical to those of the Southeast Asian compounds.

Despite the relatively central role that co-compounds play in the lexicons of many Southeast Asia languages, the heavy use of compounds of this kind, particularly those consisting of near synonyms, is often characteristic of a particular speech style. Furthermore, certain poetic styles exploit the pairings made in coordinate compounds as part of their structure. Take the following examples from Hmong (Mong Leng):

- (119) a. nqâi - tǎi
muscle skin
'flesh'

¹The cuneiform signs that have been used are actually the Akkadian forms (from a particular period). I suspect that any reader erudite enough to be bothered by this fact is also erudite enough to apply the appropriate transformations in order to arrive at the intended forms.

b. ɛ́oŋ - ntoŋ
bamboo wood/tree
'woody vegetation'

c. nqâi í daj tǎi í p^hó
muscle one piece skin one hide

ɛ́oŋ í tsɔŋ ntoŋ í tʃó
bamboo one grove wood one tree

'(We are) flesh of one cut, skin of one hide, bamboo of one grove, wood of one tree.'

(120) a. tɔŋ - hǎ
mountain valley
'hills and valleys; terrain'

b. lû tua tu nóŋ, tʃhǎ tù nó qá tɔŋ
weasel die sever seed remain CLF live bottom mountain

nà tua tu tʃǎ, tʃhǎ tù nó qá hǎ.
rodent die sever kind remain CLF live bottom valley

'Though weasels die till their seed is cut off, / there is still the one at the base of the mountain. / Though rats die till their kind is cut off, / there is still the one at the bottom of the valley.'

It is clear that the users of poetic devices of this kind are able to analyze many compounds into their consistent parts, even if the combination of those parts is somewhat lexicalized (like nqâi-tǎi 'flesh'). The ability of speakers to identify and manipulate the conjuncts of lexicalized coordinate compounds is made even more evident by the characteristics of so-called elaborate expressions². These constructions are coordinate compounds, typically consisting of four syllables, where each of the conjuncts has two internal constituents. Most often, each of the two constituents is a subordinat-

²This term was introduced by Haas (1964), with reference to Thai, and was later popularized among Southeast Asianists through the work of James Matisoff, particularly Matisoff (1973a, 1989).

ing compound. In the prototypical case, one subconstituent is repeated in each of the conjuncts (most frequently the head of the construction).

(121) chɔ̌ khô - chɔ̌ šá
 person words person spirit
 ‘One’s inner nature; one’s habits and personality’ (Lahu; Matisoff 1989, in press).

(122) tú tsaj̄ - tú ntje
 boy able boy sharp
 ‘scholars; intellectuals’ (Mong Leng)

These constructions vary in their pragmatic import. In Lahu, these expressions are most frequent in flowery and colorful speech registers (Matisoff 1973a). In Mong Leng, too, these expressions are more common in oratorical and narrative speech than in day-to-day conversation. However, even in relatively casual registers, they are not uncommon. For example tú lâŋ-tú lùam ‘boy-trade-boy-business’ is used by Mong Leng speakers in North America as the ordinary word for ‘businessman’, regardless of speech style. In other words, while elaborate expressions have some relation to the patterns created by the poetic devices described above, they must be treated as a principle facet of word structure in languages like Hmong, not mere poetic parallelism.

These constructions are interesting, among other reasons, for the fact that they involve the “intercalation” of lexical items. Potentially idiomatic coordinate compounds can be divided into parts by other lexical material without losing their idiomatic semantic content:

(123) a. ntáu - ntăi
 cloth paper
 ‘literature; education’
 b. kaj̄ ntáu - kaj̄ ntăi
 study cloth study paper
 ‘to learn to read and write, to become educated’ (Mong Leng)

Phenomenon of this type were labeled “ionization” by linguist-chemist Y.R. Chao (1948). Occasionally, words that were historically monomorphemic (or, at very least, not coordinate in their structure) have been reanalyzed as coordinate compounds and “intercalated” with some other lexical item. For example, the word *phôŋ.jj̩* ‘friend’ comes from Chinese 朋友 *péngyǒu*, and was not morphologically analyzable in Chinese at the time that it was borrowed into Hmong (probably within the last 500 years)³. However, it is sometimes treated as if it consists of two elements meaning ‘friend’ that are compounded in a coordinate structure:

- (124) a. ua p^hôŋ - ua jj̩
do friend₁ do friend₂
‘be friends’
- b. kě p^hôŋ - kě jj̩
way friend₁ way friend₂
‘friendship’

This and other evidence strongly suggests that the two members of the “couplet” (that is, the parts of each conjunct that differ from one another) may actually still form a word, despite the presence of phonological (and probably morphological) material interposed between its two parts. This same type of phenomenon is to be found in Qe-Nao, where the (apparently monomorphemic) word *s^hè.ɬä* ‘empty’, may be divided in exactly the same way as *phôŋ.jj̩* (Pan and Cao 1972):

- (125) a. tá s^hè - tá ɬä
come empty₁ come empty₂
‘to come in vain’

³On internal evidence, we can know that the source of this word was Mandarin, since it has an aspirated onset but is in the A2 tone and this results from a sound change that was confined to Mandarin. However, it cannot be from the most recent stratum of loanwords from Mandarin either; otherwise, we would expect **phôŋ.jäuu*, which is unattested. It is true that this word was originally a co-compound—transparently so in the Old Chinese period. However, it had become largely unanalyzable by the Mandarin period.

- b. ẽ s^hè - ẽ ʌã
do empty₁ do empty₂
‘to do in vain’

This leads to an interesting conclusion in the case of what have been called “pseudo-elaborate expressions.” Like elaborate expressions strictly construed, these compounds have two coordinated conjuncts, each of which has two morphological parts (and typically two syllables). In these forms, though, there is no repeated element: the structure is AB-CD, where A and C form one couplet and B and D form a second couplet. These couplets most frequently correspond to lexicalized coordinate compounds, and which may be quite idiomatic in their semantics. Their semantic contribution to the expression as a whole is *not* equivalent to that of each of the parts, but is the idiomatic meaning of the associated coordinate compound. Take, for example, the following compounds from Mong Leng:

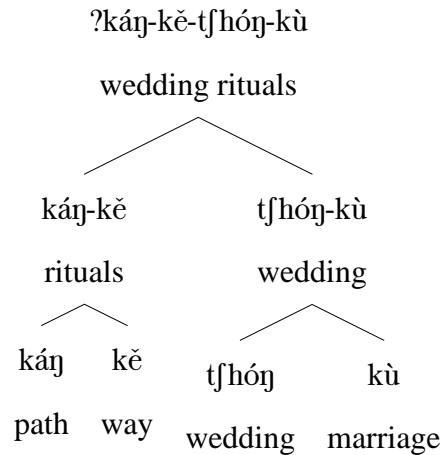
- (126) a. káŋ - kě
path way
‘ceremony; ritual’
- b. tʃhónŋ - kù
wedding marriage (bound)
‘wedding’

One has idiomatic semantics; the other includes a bound form. These may be compounded together:

- (127) káŋ tʃhónŋ - kě kù
path wedding way marriage
‘marriage rituals’

Semantically, this compound behaves as if it were composed of one cocompound, tʃhónŋ-kù ‘wedding’, modifying the other, káŋ-kě ‘ceremony; ritual’, as depicted in example (128):

(128)



Such compounds are perfectly permissible in Hmong (129):

- (129) a. nṽ - tʃĩ
official lord
'leaders; rulers'
- b. té - chàì
land place
'land; country'
- c. nṽ - tʃĩ - té - chàì
official lord land place
'rulers of the country'

However, in the case of **káj-tʃhóŋ-kě-kù**, the two compounds are (on the surface, at least) intercalated with one another to form what look (from a different point of view) like two independent subordinating compounds conjoined to form a larger coordinate compound. In some sense, two parses of the compound are possible—one which captures the head-modifier relationship between the roots in each conjunct and another that captures the lexical relationship between the words in each couplet. The importance of this set of facts will become evident later in this chapter, where we will be concerned primarily with the sequence of elements within a coordinate compound. It is sufficient at this point to note that the order of elements in an elaborate couplet is always (with one type of exception to be discussed below) the same as in the corresponding

coordinate compound.

5.1.2 The relation of co-compounds to other constructions

The co-compounds discussed here bear a certain relationship to several other types of morphological and phrasal constructions. In order to appreciate the significance of the phonological properties displayed by co-compounds in Southeast Asian languages like Jingpho and Hmong, it is important to understand how these constructions are different from (and similar to) constructions of three types: copulative compounds (or *dvandva* compounds), irreversible binomials, and a continuum of reduplicative and reduplication-like phenomena where the two conjuncts (roughly, base and reduplicant) are required to be different in a specific respect.

5.1.2.1 Copulative compounds

Many European languages, including English, allow the existence of a class of compounds with coordinate structure that have been traditionally called *copulative compounds* or *dvandva* compounds (after the term for such compounds in the Sanskrit grammatical tradition, it being a *dvandva* compound itself). In fact, compounds of this type are found quite widely in the languages of the world. Olsen (2000) makes a thorough study of their morphosyntactic properties and finds considerable internal diversity within the class of constructions that have been called *copulative compounds* or *dvandva compounds*. Their uniting characteristic is that they lack an apparent morphological or semantic head. In that respect, co-compounds in languages like Hmong and Chinese should be considered a subset of this larger class. As a result of their “headlessness,” compounds of this type may typically be reordered with no change in meaning:

(130) a. *poet-prophet* ‘person who is a poet and a prophet’

- b. *prophet-poet* ‘person who is a prophet and a poet’
- (131) a. *doctor-lawyer-architect* ‘person who is a doctor, a lawyer, and an architect’
- b. *architect-doctor-lawyer* ‘person who is an architect, a doctor, and a lawyer’

In conventionalized instances of this construction type, such as English *singer-songwriter* and *AFL-CIO* or German *Baden-Württemberg* and *Schleswig-Holstein*, the order of constituents in a copulative compound may be fixed. However, there seems to be no grammatical constraint that fixes the sequence of elements. In this respect, as will be seen in Section 5.2, the copulative compounds of English are different from the coordinate compounds of many Southeast Asian languages.

In their semantics, too, copulative compounds may differ from the coordinate compounds we have been examining. In terms of their semantics, coordinative compounds can be grouped loosely into three types: restricting, inclusive, and generalizing. Compounds like English *fighter-bomber* are restrictive: each constituent imposes an additional constraint on entities to which the compound can refer. A *fighter-bomber* must be both a fighter and a bomber; to be just a fighter or just a bomber is not enough. A compound like *Alsace-Lorraine*, in contrast, refers not to a region that is simultaneously Alsace and Lorraine, but to Alsace and Lorraine together, as a larger region. Cross-linguistically, this (inclusive type) seems to be the most common type of *dvandva* compound, and examples may be found in many languages:

- (132) a. yama - tera
 mountain temple
 ‘mountain and temple’ (Japanese; Itô and Mester 1986)
- b. aana - kuṭṛa - kaḷə
 elephant horse PL
 ‘elephants and horses’ (Malayalam; Mohanan 1986:90)
- c. candra - āditya - u
 moon sun DL
 ‘the moon and the sun’ (Sanskrit; Olsen 2000:282)

- d. $\text{na} \quad - \text{tsí}$
 woman man
 ‘husband and wife; father and mother’ (Mong Leng)
- e. $\text{ŋfi} \mid - \text{mpe} \mid - \text{mfir} \mid$
 cow pig horse
 ‘cows, pigs, and horses’ (Pa-hng; Mao and Li 1997:68)

Less common cross-linguistically, but exceptionally common in Southeast Asia, are generalizing compounds. This is an imprecise category, including synonym compounds, antonym compounds, and compounds of elements drawn from a single semantic field. The antonym compounds, in particular, are not an internally homogeneous class. For example, the semantic properties of Chinese 多少 *duō-shǎo* ‘how many? (“many-few”)’ are quite different from those of Mong Leng *já-qè* ‘high and low; everywhere (“high-low”)’ . More nuanced discussions of these topics are to be found in Chao (1968); Thompson (1965) and elsewhere.

5.1.2.2 Irreversible binomials

While the co-compounds under discussion are probably best seen as a subset of a larger class of coordinative compounds, including those of the *dvandva* type, they also resemble another type of expression: what Malkiel (1959) called “irreversible binomials”. These expressions (also called “freezes” by Cooper and Ross (1975) and others) are, to use Malkiel’s definition, “sequence[s] of two words pertaining to the same form-class, placed on an identical level of syntactic hierarchy, and ordinarily connected by some kind of lexical link” whose order is fixed by convention.

- (133) a. *warp and woof* e. *kith and kin*
 b. *weal and woe* f. *bed and board*
 c. *wild and woolly* g. *birds and bees*
 d. *rough and ready* h. *part and parcel*

- | | |
|---------------------------|-----------------------------|
| i. <i>vim and vigor</i> | p. <i>high and low</i> |
| j. <i>hale and hearty</i> | q. <i>peace and freedom</i> |
| k. <i>aid and abet</i> | r. <i>times and seasons</i> |
| l. <i>odds and ends</i> | s. <i>God and country</i> |
| m. <i>wine and dine</i> | t. <i>stars and stripes</i> |
| n. <i>hither and yon</i> | u. <i>ways and means</i> |
| o. <i>here and there</i> | |

These binomial expressions are like the Southeast Asian co-compounds in the types of semantic relationships that exist between the members (as well as a general preference for either alliteration or rhyme between the two conjuncts). In fact, in some cases, there are direct equivalences between irreversible binomials and co-compounds in languages like Hmong. Compare, for example, English *hither and yon* with Mong Leng *u-nũa* ‘hither and yon (“PROX-DIST”)’.

One salient property of these collocations is the presence of “bound” forms. Not infrequently, one—and sometimes, both—of the conjuncts in a binomial no longer occur outside of that fixed phrase. *Kith* is a classic example of a word of this type, but *weal* from *weal and woe* also fits into this category, and this example could be multiplied *ad libitum*. This is also frequently the case with co-compounds. In Jingpho, for instance, there is a compound *à.mjàʔ-fǎ.lá* ‘arrest indiscriminately’ composed of *à.mjàʔ* ‘seize’ and *fǎ.lá*, a stem of uncertain meaning that no longer occurs outside of this expression. In Mong Leng the compound *ncâu-lu* ‘mouth’, consists of two words for mouth: *ncâu*, the commonly used word, and *lu*, which no longer occurs as an independent word with this meaning (though it does occur as the classifier for utterances). Also, Mong Leng *nqâ-no* ‘food (“meat-rice”’) consists of one extremely common root, *nqâ*, and *no*, a root that now occurs only in this compound and elaborate expressions derived from it. Finally, in High Pacoh, the expression *qaqâs-qasêw* ‘terribly filthy’ contains a free stem

qaqâs ‘filthy’ and a bound stem qasêw, which historically meant ‘fearful’.

The principles that determine the order of conjuncts in such expressions have been a matter of discussion at least since Malkiel (1959). Malkiel himself pointed to a range of factors, both formal and semantic, that seemed to influence the sequence of elements in expressions of this kind. On the phonological side, he proposed that “smaller” words (words with less phonological substance) tend to be ordered before “larger” words. On the semantic side, he posited a whole range of factors including chronological priority, socially constructed priority, and relative strength of polarized traits.

In this same vein, Cooper and Ross (1975) attempted to delineate both semantic and phonological motivations for the order conjuncts in expressions of this type. On the semantic side, they propose principles like *Me First*, which states that “first conjuncts refer to those factors which describe the prototypical speaker,” *Divine* (*sacred* before *profane*), and *Plant* (*flora* before *fauna*). On the phonological front, they propose a whole set of features that characterize “place 1 elements:”

- (134) a. more syllables⁴
b. longer resonant nuclei (\bar{V})
c. more initial consonants
d. vowel containing a lower F2
e. fewer final consonants
f. a less obstruent final segment

They provide a series of examples illustrating each of these tendencies. It is significant, though, that for a few of them, such as (134d) and (134f), Cooper and Ross depend on evidence not from binomials proper, but from echo-reduplication constructions (for which, see Section 5.1.2.3 below).

⁴Note that this exactly contradicts the claims of Malkiel (1959).

In a recent corpus-based statistical study of binomial expressions in English (in which no distinction was made between irreversible and reversible binomials) Benor and Levy (2006) found that phonological criteria were relatively weak predictors of the order of conjuncts in these constructions. Semantic factors, in contrast, were found to be strong predictors of sequence. Factors of this type have also been argued to play a major role in determining the order of conjuncts in Chinese co-compounds (Chao 1968; Quan 1990; Bottero 1996). It is not clear, however, to what extent these results can be extended to the irreversible binomials and “freezes” discussed by Malkiel (1959) and Cooper and Ross (1975).

In fact, there are relationships within these binomial expressions which are clearly of a phonological nature, though not all of them pertain directly to the sequence of conjuncts. These constructions tend to shade into the reduplicative constructions to be discussed next. In some cases, one conjunct is simply a quasi-reduplicative match of the other. In *spick and span*, for instance, the *spick* part was simply an extension added to the expression *span-new*. *Spick and span*, as such, is a shortening of this expression. Interestingly, the high front vowel in *spick* and the low vowel in *span* follow the template for English echo-reduplication, as will be discussed in Section 5.1.2.3 below, and also the ordering generalization for Jingpho coordinate compounds (high before mid, mid before low), as described in Sections 5.2.2 and 5.5 below. The same may be said for *hem and haw*, where the first conjunct is original and the second conjunct is a secondary “reduplicant.”

Similar cases of quasi-reduplication exist in Biblical Hebrew. The famous (and sometimes controversial) expression תָּהוּ וְתָהוּ *tōhû wābōhû* ‘formlessness; chaos (“formlessness and [formlessness]”) consists of תָּהוּ *tōhû*, which by itself means ‘waste(land)’ or ‘formlessness’ and בָּהוּ *bōhû*, which only occurs with *tōhû*, either in a binomial or in poetic parallelism. This pattern of occurrence is reminiscent of

the distribution of elements of Hmong coordinate compounds, as discussed in Section 5.1.1 above. The second conjunct (*bōhû*) appears to have been created as a rhyming counterpart to *tōhû* in much the same way as *spick* was created as an alliterative counterpart of *span*. In this way, binomial expressions bear a strong resemblance to pseudo-reduplicative compounds. Apparently, the only principled means of drawing a distinction between these two classes of constructions is to label *binomials* those expressions with an overt conjunction and reserve the terms *pseudo-reduplication* and *echo-reduplication* for expressions with no overt conjunction (and that may thus, to some extent at least, be analyzed as single words rather than syntactic phrases).

5.1.2.3 Pseudo-reduplicative compounds and echo-reduplication constructions

English, like many other languages, has a body of binary words that are characterized by repetition, on the one hand, and a mandatory difference between parts, on the other. In one set of these words, like *teeny-tiny* and *fiddle-faddle*, the two conjuncts are identical except for the nuclear vowels of the stressed syllable. In the other, far more heterogeneous set, the two conjuncts are identical except for the initial onsets (or onset clusters). Words of this type include *namby-pamby* and *pell-mell*. This type of construction appears in many languages, but these languages are disproportionately from Eurasia.

The English case is useful, not only because it is well-studied, but also because the patterns are especially relevant to the co-compound ordering generalizations we will discuss below. I will argue that the dominant patterns in English echo-reduplication can be understood in terms of some fairly general morphological principles—principles that will be of analytic value subsequently. This claim of generality may seem surprising, since there seem to be a great many patterns of reduplication, particularly in the rhyming type. Given the presence of forms like *boogie-woogie*, *fuddy-duddy*, *ragtag*, and *mumbo-jumbo*, one might suppose that there is no generalization to be made. How-

ever, one pattern is by far the most common, and that is the pattern in which the base (used here to refer to the conjunct with the largest amount of unpredictable phonological material) occurs in the second position. The copy in the first position is identical, except that the onset of the initial syllable is replaced by /h/, as shown in (135):

- (135) a. *hugger-mugger* j. *hocus-pocus* s. *humdrum*
 b. *huncamunca* k. *higgledy-piggledy* t. *hoity-toity*
 c. *hackerty-backerty* l. *hobnob* u. *hotsy-totsy*
 d. *hubble-bubble* m. *helter-skelter* v. *hobson-jobson*
 e. *hurly-burly* n. *harem-scarem*
 f. *holus-bolus* o. *hippy-dippy* w. *heebie-jeebies*
 g. *hanky-panky* p. *hurry-scurry* x. *hurdy-gurdy*
 h. *hokey-pokey* q. *handy-dandy* y. *happy-clappy*
 i. *hodge-podge* r. *humpty-dumpty* z. *hootchie-kootchie*

It is tempting to view this as a kind of “fixed segmentism,” where part of the underlying form is “overwritten” by a fixed string of phonological material in the reduplicant. This may, indeed, be part of the explanation for this pattern. However, there must be some other principle at work. Significantly, in some of these words, it is the first conjunct (with initial /h/) that served as the historical base, with the second conjunct emerging as a result of reduplication. For example, in *hurry-scurry*, it is *hurry* that served as the original base. The contemporary English verb *scurry* has its source in this reduplication construction. The same is apparently true of *handy-dandy* and *humdrum*, *mutatis mutandis*. In fact, while in some of these forms, like *hubble-bubble*, the historical base is the second conjunct, this is by no means the situation in the majority of cases. It seems, in fact, that the origin of the two conjuncts is less important than satisfying the templatic constraints associated with the construction:

- (136) a. The first conjunct must have /h/ as the onset of its first syllable.
- b. The second conjunct must have something other than /h/ as the onset of its first syllable.
- c. The two conjuncts must be segmentally identical, otherwise.

There seem to be two other preferences, namely that the onset of the second conjunct be either a labial stop or an /sk/ cluster. Some of these expressions were not originally reduplicative at all. *Hodge-podge*, for example, comes from *hotch-potch* which is derived from *hotchpot* ‘a dish containing a mixture of many ingredients’ (< OF *hochepot*). Along the same lines, *holus-bolus* seems to have come from either the expression *whole bolus* or from Greek ὅλος βῶλος ‘whole lump’. In either case, the expressions were not originally in a coordinate structure and did not rhyme. This effect seems to be similar or identical to the “aggressive reduplication” for which Zuraw (2002) seeks to account. In a large percentage of the cases, neither conjunct appears to be derived directly from an independent lexical item.

The same issues are raised by the second type of pseudo-reduplicative compounds in English, but in a fashion that will be more directly applicable to the analysis of coordinate compound ordering effects. These have been called *ablaut combinations* by Marchand (1969), but this is a somewhat unfortunate term since these compounds do not involve ablaut in a conventional meaning of that term⁵. What does characterize them is the satisfaction of a template not entirely unlike that for the rhyming compounds of the *helter-skelter* type. Here, the template requires that the tonic vowel of the first conjunct be a high front vowel (typically /ɪ/) while the corresponding vowel of the second conjunct must be a low vowel (either /æ/ or /ɑ/, assuming Western American English pronunciation). This means that these echo-reduplicative compounds will be of one of two types. The most common type is characterized by a /ɪ-æ/ sequence:

⁵Admittedly, the results are sometimes quite similar; cf. the case of *sing:sang*.

(137) /ɪ/ — /æ/

- | | | |
|---------------------------|---------------------------|---------------------------|
| a. <i>chit-chat</i> | j. <i>fimflam</i> | s. <i>skimble-skamble</i> |
| b. <i>click-clack</i> | k. <i>jingle-jangle</i> | t. <i>slipslap</i> |
| c. <i>clitter-clatter</i> | l. <i>knick-knack</i> | u. <i>snicker-snacker</i> |
| d. <i>crinkum-crankum</i> | m. <i>mishmash</i> | v. <i>snipsnap</i> |
| e. <i>dilly-dally</i> | n. <i>pitter-patter</i> | w. <i>ticktack</i> |
| f. <i>dingle-dangle</i> | o. <i>prittle-prattle</i> | x. <i>tittle-tattle</i> |
| g. <i>drizzle-drazzle</i> | p. <i>rickety-rackety</i> | y. <i>whimwham</i> |
| h. <i>flip-flap</i> | q. <i>riffraff</i> | z. <i>zigzag</i> |
| i. <i>fiddle-faddle</i> | r. <i>shilly-shally</i> | |

The less common pattern is the one in which /ɪ/ precedes /ɑ/:

(138) /ɪ/ — /ɑ/

- | | | |
|----------------------|-----------------------|-------------------------|
| a. <i>clipclop</i> | g. <i>jiggy-joggy</i> | m. <i>ticktock</i> |
| b. <i>crisscross</i> | h. <i>pingpong</i> | n. <i>tiptop</i> |
| c. <i>dingdong</i> | i. <i>pishposh</i> | o. <i>wibble-wobble</i> |
| d. <i>dripdrop</i> | j. <i>plip-plop</i> | p. <i>wimbly-wambly</i> |
| e. <i>hiphop</i> | k. <i>singsong</i> | q. <i>wishwash</i> |
| f. <i>flipflop</i> | l. <i>slipslop</i> | r. <i>wishy-washy</i> |

And even rarer pattern has /i/ before /ɑ/:

(139) /i/ — /ɑ/

- | | |
|-------------------------|-------------------|
| a. <i>gee-gaw</i> | c. <i>see-saw</i> |
| b. <i>teeter-totter</i> | d. <i>hee-haw</i> |

- g. $\text{DU.BU.}|\text{U.GUD}| . \text{DA.BA.AL}$ 'make liquid noise'
 dubul - dabal
 𒍪
 ZA
 za
- h. $\text{PI} . \text{BA ZA}$
 wu - wa za
 'make noise'

In at least one case, the alternation is between /i/ and /a/, rather than /u/ and /a/:

- (141) $\text{ZI.IG} . \text{ZA.AK ZA}$
 zig - zag za
 'to make rumbling noise'

Reduplicative expressives of this type are quite common and it is also quite common for them to be “alliterative,” that is, for the two conjuncts to vary in their rhyme. In the Sumerian case, this alliteration is not achieved by concatenating two lexical words, nor is it the result of concatenating a word with a minimally differentiated double. Rather, it is the filling out of phonological template with onomatopoeically-motivated consonantal material.

Echo relationships are not necessarily templatic, however. In Eastern A-Hmao, there is a productive echo-reduplication construction in which the tonic vowel of the first conjunct must always be (replaced by) either /i/ or /u/ (Wang and Wang 1996). There is a principled basis for which of these segments is chosen, which will be discussed below in Chapter 6. At this point, it is sufficient to observe that the tonic vowels of the first and second conjuncts must always be different and that the conjunct of the first vowel must be high:

- (142) a. $\text{a}|\text{ndli}| - \text{a}|\text{ndliau}|$ 'margins of fields'
 RED leaf
 'leaves, flowers, etc.'
- b. $\text{a}|\text{ndi}| - \text{a}|\text{ndu}|$
 RED side
- c. $\text{a}|\text{mu}| \text{a}|\text{ma}|$
 RED eye

probably has its source in this very same constraint.

So far, we have seen two kinds of disharmony relationships in echo reduplication: non-identity of initial onsets and non-identity of the vowels in syllables bearing primary stress⁷. It seems, in fact, that there are a limited range of prominent positions to which these non-identity constraints may make reference, and the two that are most important for our purposes are:

- (144) a. Initial syllable
b. Syllable bearing primary stress

It is theoretically important that there be some means to establish a special relationship between the tonic vowels of the two conjuncts in a construction of this type. Following (Walker 2000a; Rose and Walker 2004; Hansson 2001), I will propose that long-distance phonological relationships, including long-distance relationships of this type, are established through correspondence⁸. As detailed in Chapter XYZ above, string-internal correspondence relationships are enforced by constraints requiring similar phonological constituents to be in correspondence. Earlier authors have typically conceived of this similarity in terms of featural content (for example, similarity in major class features). The innovation made here is to extend this kind of similarity relationship to structural position—to treat, for example, syllables that occupy the head positions in prosodic words as more similar (in a manner accessible to constraints of this class) than those that do not. The other, more interesting, claim made here is that constituents in prominent positions (head position, initial position, final position) are universally more “similar” than constituents not in these positions. That is to say, a constraint like $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ “head syllables are in correspondence,” and $\text{CORR-}\#\sigma \leftrightarrow \#\sigma$

⁷The Sumerian case is somewhat different since, in cases where the conjunct is more than two syllables long, all of the vowels in the conjunct are the same.

⁸It is the experience of this author that long-distance relationships of the non-phonological type may also be established through correspondence.

“PrWd-initial syllables are in correspondence,” always dominate constraints of the type $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ “non-head syllables are in correspondence,” or $\text{CORR-}\neg\#\sigma \leftrightarrow \neg\#\sigma$ “non-initial syllables are in correspondence,” if constraints of the latter type exist at all. Thus we predict that there are *no cases of long-distance correspondence between non-prominent constituents where there is not also correspondence between prominent constituents at an equal distance*. The other essential notion to review here is the idea of “trickle-down correspondence”: if two syllables are in correspondence, then their respective onsets and rimes are in correspondence; if two rhymes are in correspondence, then their respective nuclei and codas are in correspondence.

The mechanics of this system can be illustrated using the form *tomāto-tomāto* (where $\langle \grave{a} \rangle$ is /a/). Though this expression actually consists of the concatenation of two dialect forms, its similarity to the reduplication cases discussed above, and the fact that the stress is non-initial, makes it useful from the point of view of exposition. Let us begin with a constraint ranking such that $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ dominates NOCORR, which—in turn—dominates $\text{CORR-}\#\sigma \leftrightarrow \#\sigma$. Given *tomáto-tomáto* as an input, we predict that the tonic vowels will be in correspondence. In (145), this correspondence relationship is shown with arrows:

(145)

	tomáto-tomáto	$\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$	NOCORR	$\text{CORR-}\#\sigma \leftrightarrow \#\sigma$
(a)	tomáto tomáto	*!		*
(b)	tomáto tomáto └──┬── ↑	*!	*	
(c)	tomáto tomáto └──┬──┬── ↑ ↑		**!	
(d)	tomáto tomáto └──┬── ↑		*	*

Candidates like (a), in which neither initial syllables nor tonic (head) syllables are in correspondence, satisfy NOCORR at the expense of violating both of the CORR constraints. Since $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ dominates NOCORR, candidate (a) is eliminated. If only

the initial syllables of the two conjuncts are in correspondence, as in (b), only CORR- $\#\sigma \leftrightarrow \#\sigma$ is satisfied, and since it is dominated by both NOCORR and CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$, it is eliminated. If the initial syllables are in correspondence, and the head syllables are also in correspondence, (as in (c)) then both CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ and CORR- $\#\sigma \leftrightarrow \#\sigma$ are satisfied, but NOCORR is violated twice. Candidate (c) is eliminated because it is possible to satisfy CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ and only violate NOCORR once by placing only the head syllables in correspondence.

Now suppose that, in the same hierarchy, there is a constraint that requires sequential vowels in a correspondence relationship to be /ej/ and /a/, which we will call TEMPLATE⁹ Suppose, too, that this constraint dominates FAITH (faithfulness, standing in in this case for IDENT[feature] and MAX[segment]). If TEMPLATE was to dominate CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ then it would effectively block correspondence relationships in words that did not already satisfy the template, since it would be possible to satisfy both TEMPLATE and NOCORR by the simple absence of coindexation. If, however, CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ dominates TEMPLATE, then the tonic vowels of two prosodic words (here conjoined in a coordinate construction), will change to satisfy TEMPLATE. This is demonstrated in the tableau in (146), where stress is marked with the IPA notation and correspondence relationships are indicated with letter subscripts:

	tomāto tomāto	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	NOCORR	CORR- $\#\sigma \leftrightarrow \#\sigma$	TEMPLATE	FAITH
(a)	to'māto-to'māto	*!		*		
(b)	to'mā _i to-to'mā _i to		*	*	*!	
(c)	to'mā _i to-to'mā _i to		*	*		*
(d)	to _i 'mā _j to-to _i 'mā _j to		*!*		*	*
(e)	tā _i 'mā _j to-tā _i 'mā _j to		*!*			**

The observant reader will have noted that NOCORR is actually completely unnecessary

⁹As it stands, this constraint is incompletely formalized and insufficiently general. I have not sought to provide a complete formulation since the primary aim of this section is to explain how string-internal correspondence relationships function in echo-reduplication, facts that will be seen to be applicable to coordinate compound ordering effects as well.

in this grammar. As long as CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ dominates TEMPLATE, and TEMPLATE dominates FAITH and CORR- $\#\sigma \leftrightarrow \#\sigma$, only tonic vowels will be in string-internal correspondence and the template will be enforced across these vowels, as shown in the tableau in (147):

	tomāto tomāto	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	TEMPLATE	FAITH	CORR- $\#\sigma \leftrightarrow \#\sigma$
(147)	(a) to'māto-to'māto	*!			*
	(b) to'mā;to-to'mā;to		*!		*
	(c) to'mā;to-to'mä;to			*	*
	(d) to;'mā;to-to;'mä;to		*!	*	
	(e) tā;'mā;to-tä;'mä;to			**!	

From these examples, it should be clear how correspondence relationships can be established between syllables in head positions only, or between syllables in initial position. However, we have not seen an illustration of correspondence holding between initial syllables where these were not also the heads of feet. Cross-linguistically, however, it is necessary to grant both of these positions a privileged status relative to string-internal correspondence constraints.

In fact, both of these positions seem to play a role in another English echo-reduplication construction: the highly productive *shm*-reduplication process by which *David* becomes *David-Shmavid* (recently borrowed into American English from Yiddish). This process has been claimed to replace the first onset in the second conjunct (Alderete et al. 1999). However, as demonstrated by Nevins and Vaux (2003) (in a study with greater empirical depth) the situation is not so simple. One complication relevant to the current discussion is variation seen in *shm*-reduplications of words with second-syllable stress:

- (148) a. obscéne-shmobscéne ~ b. confúsion-shmonfú>sion ~
 obscéne-obshméne confú>sion-conshmú>sion

In one variant, the anti-identity requirement holds over the onset of the first syllable; in

the second syllable, it holds over the onset of the primary-stressed syllable. These same “anchor points” (see Yu 2002; Nevins and Vaux 2003) are the relevant points of comparison in the ordering effects that will be our primary focus. No coincidentally, they all relate in some way to prosodically prominent positions. Nevins and Vaux note that there are even cases where two prosodically strong positions are targeted by this process, providing the example *forbidden-shmorshidden*. In terms of the constraints we discussed in the hypothetical *tomāto-tomāto* example—and modifying our hypothetical TEMPLATE constraint appropriately—we can express the variation in the production of *shm*-reduplication for the base *forbidden* through the following typology:

- (149) a. *for**u**idden-shmor**u**idden*
 CORR-#σ ↔ #σ ≫ TEMPLATE ≫ FAITH, CORR-σ ↔ σ
- b. *for**u**idden-forshmidden*
 CORR-σ ↔ σ ≫ TEMPLATE ≫ FAITH, CORR-#σ ↔ #σ
- c. *for**u**idden-shmorshmidden*
 CORR-σ ↔ σ, CORR-#σ ↔ #σ ≫ TEMPLATE ≫ FAITH

One can also imagine the possibility that, if the ranking of the constraints that govern the position of stress were relatively low when compared to the CORR constraint, the position or structure of feet could change in order to satisfy both CORR-σ ↔ σ and CORR-#σ ↔ #σ with a single correspondence relationship. This variant can also be seen, in forms like *óbscene-shmóbscene*.

One aspect of *shm*-reduplication that we have not discussed thus far is the anti-identity requirement. It has often been noted that forms like **schmaltz-shmaltz* or **schmooze-shmooze* are not acceptable as instances of this construction: non-identity between the two conjuncts is required. The non-identity may be achieved, in many cases, by substituting other segmental material for the *shm* sequence. Nevins and Vaux (2003) give the following list of examples for *schmooze*:

- (150) *shnooze, flooze, shpooze, shlooze, vlooze, shplooze, shmămooze, mooze, wooze, commooze, my ass*

Thus, strategies for achieving non-identity include changing the place of articulation of either the /ʃ/ or the /m/, changing the manner of the /m/, deleting the /ʃ/, epenthesizing a vowel, substituting a different morphosyntactic construction (*my ass*, for example), and so on.

This kind of requirement is quite common and can be found in many languages. One case that is instructive for our purposes is that of Bunu (another Hmongic language closely allied to the Western Hmongic group; see 5.1), where there is a verbal reduplication process carrying the semantics, “to X casually,” where X is the simple verb. Examples of this construction are given below (Mao et al. 1982:94):

- | | |
|--|---|
| <p>(151) a. ntoɿ - ntɿuɿ
 chop RED
 ‘chop (casually)’</p> <p> b. nɿŋɿ - nɿuɿ
 wear RED
 ‘wear (casually)’</p> | <p>(152) a. ntɿuɿ - ntɿoŋɿ
 hit RED
 ‘hit (casually)’</p> <p> b. nthɿuɿ - nthɿaŋɿ
 pluck RED
 ‘pluck (casually)’</p> |
|--|---|

In this construction, the rhyme of a monosyllabic root is overwritten in the reduplicant with the diphthong /ɿu/, as seen in (151), *except* when the underlying rhyme is /ɿu/, as in (152). In these exceptional cases, the reduplicant has as its rime either /oŋ/ or /aŋ/. At some level, it seems that echo reduplication in general, like this particular case, is driven by a requirement for difference between the two conjuncts. At some level, the co-compounding of synonyms bears a resemblance to these sorts of effects in reduplication: they enforce semantic similarity but formal difference.

In Bunu, the difference between the two conjuncts is achieved through melodic overwriting, with a suppletive form of the fixed segmental material appearing when the default string, /ɿu/, would not result in non-identity. The effect in A-Hmao is similar.

In English echo-reduplication and Sumerian expressives, the difference is achieved by imposing a template. The fixed-segment-with-suppletion type of reduplication can potentially apply to any word in the language, but does not guarantee the existence of a particular relationship between the loci of comparison within the two conjuncts. The templatic type is able to enforce such a relationship—high vowel before low vowel, for instance—but cannot apply to every word in the language. Indeed, in many cases of templatic “reduplication,” there is no underlying base at all: they are reduplication only in the sense that they contain repeated segmental material. However, some languages have echo reduplication constructions that satisfy both conditions.

Jingpho (a Tibeto-Burman language of Burma, China, and India) has an echo reduplication construction that may apply to any word, but which always guarantees that the tonic vowels will be in a non-low/low sequence (Dai 1990a,c). If the tonic vowel (always the vowel in the mono-syllabic root and therefore the last vowel in a stem) of base form is non-low, the reduplicant will have, in its place, the low vowel /a/, as shown in (153). In these cases, the reduplicant appears *after* the base. If, however, the tonic vowel of the base is the low vowel /a/, as shown in (154) the reduplicant has as its tonic vowel /o/ and is ordered *before* the base.

- (153) a. mət̚sùʔ - mət̚sàʔ
 disorderly RED
 ‘chaotic’
- b. sət̚ùp - sət̚àp
 muggy RED
 ‘sultry’
- c. kəm̚jɪn - kəm̚jɪn
 wrinkle RED
 ‘in a very wrinkled state’

- (154) a. kàlɔʔ - kàláʔ
 RED quarrel
 ‘quarrel’
- b. ləphɔ - ləphɔ
 RED fallen leaves
 ‘every kind of fallen leaves’

These two apparently conflicting “goals” can be achieved by allowing the base and reduplicant (really the faithful copy and the modified copy) to be freely ordered with respect to one another. We will see later that Jingpho coordinate compounds pattern in the same way as this echo-reduplication construction. The order of the conjuncts is driven by a requirement that the tonic vowel of the first conjunct be at least as high as the tonic vowel of the second conjunct. It is even possible to view echo reduplication in Jingpho as a kind of coordinate compounding where a suitable double is not available, so the grammar produces a new double by minimally altering the original form. So similar are these two types of constructions that Dai (1990a) treats them as subtly differentiated instances of the same thing.

5.2 Survey of Co-compound ordering effects

In a great many cases, phonological factors influence the ordering of conjuncts in coordinate compounds. These types of influence may be divided, roughly, into two groups: “soft” ordering effects that are driven primarily by stylistic factors and “hard” ordering effects that are enforced by the grammar. We are concerned here primarily with the hard ordering effects, providing, as they do, a greater insight into the architecture of the grammar than the soft effects. However, it is necessary to explore the soft effects briefly, both in order to contrast them with the grammatical effects and to determine what light they can shed upon this class of phenomena generally.

5.2.1 Soft (stylistic) ordering effects

Numerous languages seem to have optional, phonologically-driven stylistic effects that influence the order of morphological and syntactic objects. An example of this—though its status as a grammatical or stylistic phenomenon is still debated—is Heavy NP Shift (Ross 1967), the process by which large, non-subject noun phrases are post-

posed to the right edge of a sentence:

- (155) a. I gave [the book] to Rahel.
b. ?? I gave to Rahel [the book].
- (156) a. ? I gave [the book that reveals the dark secrets of our people's history in lurid detail] to Rahel.
b. I gave to Rahel [the book that reveals the dark secrets of our people's history in lurid detail].

The exact nature of this phenomenon has been debated. However, recent psycholinguistic experimentation suggests that the driving force behind this process is the phonological weight of the NP potentially being “moved” (Stallings et al. 1998).

It is not surprising that similar constraints enter into the ordering of constituents in coordinate constructions. These constraints may be quantitative, as has been suggested for English; may be driven by tone, as in Chinese; or may be based on vowel quality, as is the case in Lahu. A brief survey of these cases follows. A more detailed treatment of the Chinese and Lahu cases will be provided below.

English Various scholars, like Malkiel (1959) and Cooper and Ross (1975), have proposed that there are phonological constraints that influence the sequence of conjuncts in irreversible binomials. Malkiel identified a “small-before-large” condition. Cooper and Ross explored a number of additional conditions. Specifically, they argued that the tendency to appear first in a sequence of this kind was strongly influenced by the relative F2 of vowels and the “obstruence” of final consonants. In this vein, *dribs and drabs*, *spic and span*, *hem and haw*, and *(by) guess or gosh* appear to have their characteristic orders because of constraints on vowel quality. Likewise, the linear organization of binomials like *(by) hook or crook*, *kith and kin*, and *bread and water* may be influenced by consonantal features.

Chinese Ting (1969, 1975) observed that the sequence of conjuncts in many Chinese coordinate compounds was predictable based upon the tone borne by each of the conjuncts in Middle Chinese. He found this to be true not only of modern Mandarin co-compounds (Ting 1969) but also of co-compounds from texts of the Old Chinese period, long before Middle Chinese (Ting 1975). Thus, there are many compounds like 天地 *tiāndì* ‘heaven-earth’, with the Middle Chinese tones 上去 *shang-qu* ‘rising-departing’, and relatively few compounds like *地天 *dì-tiān*, with the Chinese tones ‘departing-rising’. The whole set of tendencies can be reduced to a scale: 平 *píng* < 上 *shàng* < 去 *qù* < 入 *rù*, or *level* < *rising* < *departing* < *entering*. He found that, for Old Chinese, his model could predict the correct form for about 80% of the data. More recent investigations have led to similar results (Bottero 1996). These effects seem to be rather strong tendencies, but do not look like categorical grammatical effects. This case will be discussed in greater detail below, along with a possible account of the origin of the pattern.

Lahu Lahu, a Tibeto-Burman language of China, Burma, Thailand, and Laos, has interesting vowel-quality-related ordering effects in coordinate compounds, specifically in elaborate expressions (Li 2004). This effect, however, is only a strong tendency and the number of exceptions is large. In general, conjuncts containing the vowel /o/ tend to be ordered first. Lacking such a conjunct, those containing the high vowels /i i u/ tend to be in the first position. If none of these are available, a conjunct containing non-front mid vowels /ə ɔ/ occurs in the first position. Barring this, conjuncts containing front mid vowels /e ɛ/ are ordered first. Conjuncts containing /a/ are ordered last in the great majority of cases. Assuming this set of precedence relationships, we can predict the correct order of Lahu co-compounds for about 68% of the data, on the basis of vowels alone. Tones also have a significant influence upon the order of conjuncts in co-compounds.

This case is discussed in greater detail below.

These soft ordering effects seem to originate in a variety of ways. Some may have started out as grammatical constraints that were later undermined by changes in the phonological system, resulting in the existence of non-categorical ordering generalizations within the lexicon. The preponderance of prototypes still displaying the phonologically driven order may favor the formation of new compounds with the same sequence, so that the pattern has some productivity without having the status of a full grammatical constraint. The Lahu and Chinese cases are more than likely of this type. In other cases, though, including that of English, there is no evidence that the ordering effects that can be observed in irreversible binomials or other coordinate structures were ever parts of the grammar in the same sense that the rules that select allomorphs for the English plural morpheme are (to name but one example).

5.2.2 Hard (grammatical) ordering effects

However, there are ordering effects that are largely categorical. These effects take many of the same forms as the soft ordering effects described above. These effects are superficially similar to those non-categorical ordering generalizations: they may make reference to tone and vowel quality; and when they refer to vowel quality, they tend to result in high vowels being ordered before mid or low vowels (resulting in sequences similar to those seen in English, A-Hmao, and Sumerian echo reduplication). They differ in the sense that they characterize almost all of the coordinate compounds in the lexicon (or all of a principled subset of the co-compounds, as is the case in Mong Leng). Four examples are briefly described below. More detailed analyses will be given for each of these languages.

Jingpho In Jingpho, a Tibeto-Burman language, the order of coordinate compounds can be predicted based upon the relative heights of the tonic vowels in each of

the conjuncts. Constituents are always ordered such that the higher tonic vowel precedes the lower vowel.

Tangkhul Tangkhul, another Tibeto-Burman language, appears to have an ordering effect very similar to that of Jingpho.

Hmong In Hmong, a Western Hmongic language, the order of coordinate compounds is predictable on the basis of tone. Unlike the Chinese example, this effect is incredibly robust. As will be discussed below, over 97% of the compounds in a corpus of 168 items can be predicted by a model in which the eight tones are placed in a hierarchy such that falling precedes high, high precedes creaky, creaky precedes low, low precedes breathy, breathy precedes rising, rising precedes a second breathy tone, and this breathy tone precedes mid. In contrast to the Jingpho and Tangkhul cases, this scale does not track some natural phonetic dimension. Like the Lahu¹⁰ and Chinese cases, it displays a degree of phonetic arbitrariness. Unlike these cases, though, Hmongic conjunct ordering is a categorical grammatical process rather than a stylistic preference that interacts with a variety of other factors in determining the order of conjuncts. This case, using data from the Mong Leng dialect, will be analyzed in great detail below.

Qe-Nao Qe-Nao, an Eastern Hmongic language¹¹ is like Hmong in displaying a tone-driven co-compound ordering process. Evidence suggests that this process and the one in Hmong share a common origin. From the standpoint of abstract struc-

¹⁰The Lahu vowel-ordering effect shows a certain amount of phonetic arbitrariness. The weaker tone-ordering effect, however, is transparently based upon phonetics.

¹¹Both Hmong and Qe-Nao belong to a high-level subgroup within Hmongic that I call “Southern Hmongic.” However, the relationship between these two languages is quite distant. Qe-Nao is much more distantly related to both Hmong and Qe-Nao.

ture, they look quite similar. However, from the standpoint of phonetic substance, they are not at all the same.

Qo-Siong In the Northern Hmongic language Qo-Siong, as in Hmong and Qe-Nao, co-compounds are ordered with reference to tone. However, evidence suggests that this phenomenon has a different origin than the effects seen in its southern cousins. Like the effects in Hmong and Qe-Nao, however, the ordering generalization in Qo-Siong is, to some degree phonetically, arbitrary.

5.3 Co-Compounds in Chinese

Of the cases of ordering effects that apply specifically to coordinate compounds, the most widely discussed has been that of Chinese (and specifically Old and Middle Chinese). The widespread knowledge of this phenomenon (among Asianists, at least; see e.g. Lien 1989), particularly the knowledge that it is non-deterministic, might lead those who did not know better to assume that all phonologically-motivated ordering effects were of this type. It would be worthwhile to examine this case just to establish how it is different from the examples of “hard” ordering constraints that will be discussed below. Beyond that, though, this example may show (tentatively) an example of one mechanism by which an ordering constraint might develop, namely, through the avoidance of undesirable sequences of sounds word-medially.

As was mentioned above, the ordering generalization in Chinese is usually described (and is probably best described) in terms of the four Middle Chinese tone categories. While it is traditional to refer to each of these categories as a tone (聲調 *shēngdiào*) it is more accurate to say that three of them represent contrasting tones while the fourth represents a syllable-type in which there were no tonal contrasts (syllables ending in an oral stop). Though there has been much speculation and debate on the subject, little is known about the phonetic values of these categories, so they are

typically referred to by linguists according to their traditional (exemplary) labels: 平 *píng*, 上 *shǎng*, 去 *qù*, and 入 *rù*, which are usually rendered in English as *level*, *rising*, *departing*, and *entering*. Syllables with the entering “tone” are those with an oral stop coda: /p/ /t/ or /k/. The observation of a number of investigators has led to the discovery that the order of conjuncts in co-compounds gravitates towards a transitive scale of linear precedence relationships:

(157) 平 *level* < 上 *rising* < 去 *departing* < 入 *entering*

The common sequences are illustrated here (with data taken from Ting 1975 and Bottero 1996:

- | | |
|---|---|
| (158) a. 多 <i>duō</i> 少 <i>shǎo</i>
平 <i>píng</i> 上 <i>shàng</i>
many few | e. 買 <i>mǎi</i> 賣 <i>mài</i>
上 <i>shàng</i> 去 <i>qù</i>
buy sell |
| b. 天 <i>tiān</i> 地 <i>dì</i>
平 <i>píng</i> 去 <i>qù</i>
heaven earth | f. 酒 <i>jiǔ</i> 食 <i>shí</i>
上 <i>shàng</i> 入 <i>rù</i>
liquor food |
| c. 衣 <i>yī</i> 服 <i>fú</i>
平 <i>píng</i> 入 <i>rù</i>
clothes clothes | g. 善 <i>shàn</i> 惡 <i>è</i>
上 <i>qù</i> 入 <i>rù</i>
good evil |
| d. 老 <i>lǎo</i> 幼 <i>yòu</i>
上 <i>shàng</i> 去 <i>qù</i>
old young | h. 抱 <i>bào</i> 括 <i>kuò</i>
去 <i>qù</i> 入 <i>rù</i>
enwrap embrace |

Ting (1975) found that for approximately 80% of the co-compounds in Classical Chinese texts, the sequence of constituents could be determined on tonal grounds. The *píng* < *shǎng* < *qù* < *rù* model was found to predict correct linear sequences with similar degrees of reliability in different domains by Chen and Yu (1979) and Bottero (1996). Examining all of the co-compounds in a list of 3000 words frequently used in Modern Standard Chinese (普通话 *pǔtōnghuà*), Chen and Yu found that the order of 80% of co-compounds could be predicted based on the Middle Chinese tones

for the lexical items in question and 79% (of the cases where there was a tonal difference between the two constituents) could be predicted on the basis of the Mandarin (*pǔtōnghuà*) tones. Looking only at antonymous co-compounds in classical and early modern texts, Bottero found that approximately 78% of these were phonologically predictable in their internal ordering. There are two findings that seem to be consistent across these studies: A large majority of Chinese co-compounds follow the tone-ordering generalization, across a rather vast stretch of history (about 3000 years). However, there is no evidence that this “rule” was ever categorical in historical times.

Ting (1975) took the fact that this ordering generalization stretched back to the Old Chinese period as evidence that Old Chinese was already tonal, contra the claims of scholars like Haudricourt (1954b,a), Pulleyblank (1962, 1973), and Mei (1970), who proposed that Middle Chinese tones arose from Old Chinese coda consonants of various manners. However, as Bottero (1996) points out (in an insight he attributes to Laurent Sagart), it may be easier to account for the ordering generalization in terms of these final consonants than in terms of tones. In the proposal that originated with Haudricourt (1954b,a), and which is now widely accepted among Sinologists (see, e.g. Baxter 1992; Sagart 1999 but see also Norman 1988) these sources of the tones were as follows (where P represents ‘plosive’):

- (159) CV, CVN > 平 *píng* ‘level’
 CV?, CVN? > 上 *shǎng* ‘rising’
 CVs, CVNs > 去 *qù* ‘departing’
 CVP > 入 *rù* ‘entering’

Bottero suggests that the ordering effect may result from avoidance of certain types of consonant clusters within a word. By ordering stems in the *píng* category first, the creation of obstruent-consonant clusters is minimized. Likewise, by preferentially ordering stems in the *shǎng* category before those in the *qù* and *rù* categories, buccal-

obstruent-consonant clusters are minimized. In the same way, always ordering *rù* category stems last minimizes the possibility of creating plosive-consonant clusters.

We could easily state this generalization in Optimality Theoretic terms, by proposing a hierarchy of constraints against consonant clusters of various types. In one formulation, we could employ three such constraints: *-OC- (no sequences of an obstruent followed by any consonant), *-BC- (no sequences of a buccal obstruent followed by any consonant), and *-PC- (no sequences of a plosive followed by any consonant). These constraints are in a stringency relationship of the sort explored by de Lacy (2002a). That is to say, any candidate that violates *-PC- will also violate *-BC- and *-OC-, and any candidate that violates *-BC- will also violate *-OC-. Because of this property, the relative rankings of these constraints is irrelevant (except insofar as other constraints could be ranked between them).

- (160) a. *Píng* is ordered before *shǎng* in order to avoid a medial cluster of glottal stop plus consonant.

	*/mraʔ, krja/	*-OC-	*-BC-	*-PC-
☞ (a)	車馬 [krja-mraʔ]			
(b)	馬車 [mrəp-krja]	*!		

- b. *Píng* is ordered before *qù* to avoid a medial cluster of /s/ plus a consonant.

	*/sjaws, ngjan/	*-OC-	*-BC-	*-PC-
(a)	笑言 [sjaws-ngjan]	*!	*	
☞ (b)	言笑 [ngjan-sjaws]			

- c. *Píng* is ordered before *rù* to avoid a medial cluster of a plosive plus a consonant.

	*/tsoŋ, dzok/	*-OC-	*-BC-	*-PC-
☞ (a)	宗族 [tsoŋ-dzok]			
(b)	族宗 [dzok-tsoŋ]	*!	*	*

- d. *Shǎng* is ordered before *qù* because medial clusters of /s/ plus consonant are worse than medial clusters of glottal stop plus consonant.

*/lrjajs, hlaʔ/		*-OC-	*-BC-	*-PC-
☞ (a)	土地 [hlaʔ-lrjajs]	*		
(b)	地土 [lrjajs-hlaʔ]	*	*!	

- e. *Shǎng* is ordered before *rù* because medial clusters of plosive plus consonant are worse than medial clusters of glottal stop plus consonant.

*/C-rjat, hmijʔ/		*-OC-	*-BC-	*-PC-
(a)	烈火 [Crjat-hmijʔ]	*	*!	*
☞ (b)	火烈 [hmijʔ-Crjat]	*		

- f. *Qù* is ordered before *rù* because medial clusters of plosive plus consonants are worse than medial clusters of /s/ plus consonant.

*/baws, g ^w at/		*-OC-	*-BC-	*-PC-
(a)	括抱 [g ^w at-baws]	*	*	*!
☞ (b)	抱括 [baws-g ^w at]	*	*	

Such a grammar can generate the dominant pattern seen in Old Chinese. In proposing such a solution, however, we are implicitly assuming that the ordering generalization seen in Chinese is grammatical (like the Jingpho and Hmong effects that will be discussed below) rather than purely stylistic (as seems to be the case with phonological sequencing constraints on English binomial expressions). This could be seen as justified in light of the fact that this or an equivalent grammar characterizes approximately 80% of the co-compounds in the lexicon of Chinese. On the other hand, there is no principle that defines this subset of the lexicon, or the subset that does not follow the generalization.

Another problem, more interesting for current purposes, is the fact that the Old Chinese analysis functioned without resorting to representational scales. It does indeed

involve a scale—in the sense which the term is used by de Lacy (2002a)—but this is a markedness scale, roughly:

(161) a. -PC- > -BC- > -OC- > -CC-

However, this analysis only works for Old Chinese. All authorities are agreed that by the late Middle Chinese period, final -s and -ʔ had disappeared, having been completely replaced by tonal contrasts (Pulleyblank 1962, 1973; Mei 1970; Baxter 1992). At this point, the analysis given in terms of phonotactics can no longer be invoked since its motivations have disappeared. However, the ordering generalization does not seem to have surrendered to oblivion as quiescently as -s and -ʔ. Indeed, as shown by Chen and Yu (1979), this pattern has been perpetuated up to the present, influencing the formation of compounds that did not exist in the Old Chinese period. Whether this pattern is grammatical or stylistic, it relies on phonological knowledge that cannot be stated simply in terms of substantively-oriented well-formedness constraints of the type I have invoked for Old Chinese.

5.4 Co-Compounds in Lahu

Like most of the languages of Peninsular Southeast Asia, the Lolo-Burmese (and thus, Tibeto-Burman) language Lahu has a rich lexicon of co-compounds, and particularly of elaborate expressions. Take the following examples from Matisoff (1989):

- | | |
|---|---|
| <p>(162) a. ti - ve - qha - ve
 plant NOM sow NOM
 ‘plant and sow; do the plant-
 ing’</p> | <p>c. chi - cə - qô - cə
 10 kinds 9 kinds
 ‘all kinds of’</p> |
| <p>b. phí - cûʔ - nà - cûʔ
 spirit lie spirit lie
 ‘lying spirits; temptations by
 the spirits’</p> | <p>d. kù - kò - kù -
 all persons all
 khî
 appearances
 ‘everybody; each and house-
 hold’</p> |

Phonological co-compound ordering in Lahu, like Chinese, is a non-categorical phenomenon. It is clear, on the one hand, that many co-compounds in Lahu follow a certain phonological pattern, as noted by Li (2004). Li notes that there is a general tendency for the constituent of a co-compound with the highest vowel to occur first. Additionally, he suggests that there are semantic constraints on conjunct-ordering, such that the organization of many or most Lahu co-compounds can be explained on either formal or functional grounds. Li, however, does not attempt to quantify either of these predictors, leaving it unclear what their relative or absolute strengths are.

In order to resolve this question, I performed a computational analysis of a large set of Lahu data. The source data was a collection of elaborate expressions extracted from an electronic version of *The Dictionary of Lahu* (Matisoff 1989), originally used to create an appendix to an English-Lahu lexicon (Matisoff in press). This corpus of 2373 elaborate expressions appears to be the largest consolidated collection of constructions of this type from a single language. From this corpus, a set of 1005 unique “elaborate couplets” (referring to the pair of syllables within an elaborate expression which are different from one another, which typically occur in multiple elaborate expressions, and which often exist as independent co-compounds), was extracted and analyzed using a set of text-manipulation programs¹². In the analysis, phonological constituents in the first syllable of the couplet were compared with their correspondents in the second syllable. The complete set of these comparisons was compiled (on separate runs for onsets, rimes, and tones). For each of these classes of segments (onsets, rimes, and tones) computational methods were used to find the “best” scale—the total ordering of the set of types contained within a class that is “true” of the largest proportion of the comparisons in the data set¹³. A scale *S* is considered for our purposes to be “true”

¹²All of these programs were written by the author in the Python programming language and were executed on either Linux workstations or a Macintosh G4.

¹³In examining the onsets, only broad classes were employed. The reasons for this are quite

of a comparison (α, β) if β appears in the S no earlier than α . The best scale was discovered by iterating through all of the possible permutations of elements within the set (of onsets, rhymes, or tones) and calculating the proportion of comparisons of which the resulting scale (list of permuted elements) was true.

Regarding the vowels/rimes, it was found that Li's claims were correct (or nearly so). Vowel quality proves to be a fairly powerful predictor of ordering in Lahu compounds. The best scale over the vowels is true of 68% of the couplets (versus the 50% that would be expected by chance). Li was also correct to suggest that vowel height played a role, since—in the best scales—the high vowels tend towards the beginning of the scale and the low vowel is invariably at the end. However, there are a couple of interesting exceptions to this generalization. First, the vowel which is by far most likely to occur in first position is /o/. Also interesting is the fact that all of the non-front mid vowels appear before the front vowels. Thus, /ɔ/ definitely occurs before /e/ even though /e/ is higher in the vowel space. The whole set of data arranged according to the “best” scale¹⁴ is given in (163). Vowels appearing in the first conjunct are given in rows and those occurring in the second conjunct, in columns. A number in a cell represents the percentage of occurrences of those two vowels in which they are ordered as indicated (vowel in row versus vowel in column). Percentages above 50% are in boldface type:

simple: in order to find all of the best scales, it is necessary to iterate over all permutations of the elements in each class. Since Lahu has 24 consonants, this would involve examining 15,511,210,043,330,985,984,000,000 cases, which is an impossibility given current computational capacities.

¹⁴There are actually two other scales which are true of exactly the same number of couplets. These are $\{o < u < \partial < i < \dot{i} < \partial < e < \varepsilon < a\}$ and $\{o < u < i < \partial < \dot{i} < \partial < e < \varepsilon < a\}$.

(163) **Frequencies of vowel sequences in Lahu elaborate couplets**

	o	u	i	ɨ	ə	ɔ	e	ɛ	a
o	100%	77%	82%	73%	100%	67%	50%	65%	82%
u	23%	100%	74%	46%	67%	60%	66%	86%	76%
i	18%	26%	100%	63%	58%	55%	58%	65%	59%
ɨ	27%	54%	38%	100%	50%	73%	50%	69%	81%
ə	0%	33%	42%	50%	100%	58%	38%	65%	78%
ɔ	33%	40%	45%	27%	42%	100%	70%	51%	66%
e	50%	34%	42%	50%	63%	30%	100%	94%	69%
ɛ	35%	14%	35%	31%	35%	49%	6%	100%	64%
a	18%	24%	41%	19%	22%	34%	31%	36%	100%

What is immediately evident, both from the fact that the best model can account for 68% of the data and from the relatively even spread of comparisons over the possibility-space, is that the process that produced the undeniable statistical skewing of vowel distributions within Lahu co-compounds is either not a categorical, grammatical process or is no longer productive. Nevertheless, it is next-to-impossible to believe that these biases are accidental. Take the table in (164), where the probabilities of the observed frequencies of vowel patterns in Lahu elaborate expressions is given (for vowel pairs that occur more frequently than their mirror images), assuming a binomial distribution. Figures that are less than 0.05 are considered to be statistically significant here and are given in boldface type:

(164) **Probabilities of observed vowel distributions in Lahu elaborate expressions**

	o	u	i	ɨ	ə	ɔ	e	ɛ	a
o		$3.5e^{-2}$	$2.8e^{-5}$	$2.5e^{-3}$	$6.1e^{-5}$	$1.8e^{-2}$		$5.8e^{-2}$	$4.1e^{-8}$
u			$1.7e^{-5}$		$2.3e^{-1}$	$5.0e^{-2}$	$1.6e^{-2}$	$2.0e^{-8}$	$2.9e^{-10}$
i				$7.8e^{-2}$	$1.9e^{-1}$	$7.7e^{-2}$	$1.9e^{-1}$	$1.4e^{-2}$	$3.7e^{-3}$
ɨ		$9.7e^{-2}$				$6.3e^{-5}$		$5.8e^{-3}$	
ə						$1.2e^{-1}$		$5.8e^{-2}$	$4.5e^{-8}$
ɔ							$2.0e^{-3}$	$9.4e^{-2}$	
e					$2.2e^{-1}$			$2.4e^{-4}$	$2.1e^{-4}$
ɛ									$1.2e^{-3}$
a									

This shows that, out of a total of 45 cells consistent with our scale model, 20 correspond to relationships that are statistically significant ($p < 0.5$). If we take the scale as a whole, the probability of the 1005 couplets following the scale in almost 685 cases by chance alone is less than 1.0×10^{-5} ; however, this number is deceptive, since there is no motivation for the scale outside of the distribution of vowel combinations in elaborate expressions and the scale is devised specifically to maximize the probability of the data relative to the linear precedent statements implicit in it. Thus, reporting this statistic would amount to a kind of data fishing.

What Li did not address is the fact that there are also statistical tendencies regarding the distribution of tones in Lahu co-compounds. Like the vowel-quality effects, these biases are too strong to be explained by chance alone (at least for some combinations). The probabilities that the apparent biases for particular ordering of tones are the result of chance are given in (165):

(165) **Probabilities of tonal co-occurrence patterns in Lahu co-compounds**

	$\acute{x}?$	\bar{x}	\grave{x}	x	\acute{x}	\hat{x}	$\hat{x}?$
$\acute{x}?$		$3.8e^{-1}$		$1.3e^{-1}$	$1.8e^{-2}$	$1.4e^{-3}$	$1.1e^{-3}$
\bar{x}			$1.7e^{-1}$	$3.0e^{-3}$	$6.8e^{-4}$	$6.9e^{-4}$	$5.8e^{-2}$
\grave{x}				$5.9e^{-3}$	$1.1e^{-1}$	$1.7e^{-2}$	$1.7e^{-2}$
x					$4.8e^{-2}$	$8.2e^{-2}$	$1.2e^{-1}$
\acute{x}						$9.6e^{-2}$	
\hat{x}							$8.1e^{-2}$
$\hat{x}?$					$2.5e^{-1}$		

The scale that best describes this distribution, accounting for 63.3% of the data, is of a very interesting nature:

(166) low-checked \prec very low \prec low-falling \prec mid \prec high-rising \prec high-falling \prec high-checked

Translating this scale into Chao tone numbers, following (Matisoff 1973a), the pattern becomes even more clear:

(167) 2q \prec 11 \prec 21 \prec 33 \prec 45 \prec 54 \prec 4q

The starting pitch of the non-checked tones increases monotonically across the scale. There appears to be a certain logic to this. Tibeto-Burman languages, like many languages of Southeast Asia, tend to assign greater prominence to syllables at the right edge of prosodic constituents. Syllables at the left edge of compound words tend to be destressed and reduced, leading to the formation of what are now widely called *sesqui-syllables*, following Matisoff (1973b). It makes intuitive sense that higher tones—which should be more prominent than lower tones—should be preferred in prominent positions. This intuition has been formalized in various places, notably (in reference

to the current case), by de Lacy (2002b). While de Lacy's proposal and argument are flawed in numerous respects, and are perhaps not defensible in shaping our view of phonological universals, it seems quite likely that they do reflect the kind of metalinguistic awareness that could guide stylistic decisions. There is no principled reason that stylistic decisions, too, could be grammaticalized in the shared domain of phonology and morphology, just as related patterns of usage can be seen to effect the development of morphosyntax. It seems likely that the hard ordering effects we will discuss next—both those that are phonetically transparent and those that are phonetically opaque—began as the kind of substantively natural tendency that is exemplified by the Lahu tone scale. The difference, in the languages with hard ordering effects, is that the proportion of the co-compounds in the lexicon that followed a systematic tendency came to be so large that learners reinterpreted the statistical skewing as the result of a grammatical rule rather than a shared stylistic preference.

5.5 Co-Compounds in Jingpho

In the cases that we have examined up to this point, the influence of phonology upon the ordering of coordinate compounds has been relatively weak. Certainly, there is a preference, in English, Lahu, and especially Chinese, for coordinate structures to be ordered according to some phonological rationale. In the case of Chinese, at least, this influence may have had a general phonotactic motivation, but none of these relationships seem appropriate to be described in terms of the formal grammar. They appear to take the form, largely, of stylistic effects. The case of Jingpho (a Tibeto-Burman language of Burma, China, and India) is quite different. In Jingpho, the sort of effect that is merely an inclination in English and Lahu is a grammatical imperative.

5.5.1 Ordering effects

It was noted by Dai (1990a,c) that coordinate compounds in Jingpho always follow a particular schema. If the tonic vowels (which are the vowels in the roots, and the final vowels in each stem) of the two conjuncts differ in height, the stem with the higher vowel will be the first conjunct and that with the lower vowel will be the second conjunct. This is a pattern which, by now, should be familiar. A similar sort of high-low pattern was seen in English, Sumerian, A-Hmao, and Jingpho echo reduplication (to a lesser extent, in Khmer echo reduplication as well), and in the ordering tendencies in Lahu coordinate compounds (*modulo*, of course, Lahu's unexpected tendency to put conjuncts with the vowel /o/ in first position).

In roots, Jingpho has the five vowels /i u e o a/; an additional vowel, /ə/, occurs only in (unstressed) affixes. According to Dai (1990a) and Dai and Xu (1992), vowels in co-compounds display the following sequencing possibilities:

(168) **Co-compound sequencing possibilities in Jingpho**

V1\V2	i	u	e	o	a
i	X	X	X	X	X
u	X	X	X	X	X
e			X	X	X
o			X	X	X
a					X

The only relevant factor seems to be vowel height—frontness or backness do not play a role (Dai 1990a). Thus, there are many co-compounds like:

- (169) **lùʔ - fǎ**
 drink eat
 ‘food’

but no (or almost no) co-compounds like:

- (170) * fǎ - lǔ?
 eat drink
 ‘food’ (intended)

Dai notes that there are a few exceptions to the pattern, but that they are very rare. This need not be a matter of great concern: very few morphological patterns of interest are entirely without exception. Indeed, morphologically-specific phonology is, in and of itself, a kind of exceptionality. This sequencing pattern of overwhelming strength cries for explanation in the same way that other phenomena in the morphology-phonology interface do.

5.5.1.1 Data

Examples of compounds demonstrating this generalization, all from Dai (1990a) are given below:

- | | |
|--|--|
| <p>(171) high < high</p> <p>a. mǎní - maní
 yesterday tomorrow
 “yesterday and tomorrow”</p> <p>b. mǎsìn - sǎlum
 heart heart
 “heart”</p> <p>c. kǎkhûm - ñkjin
 pumpkin cucumber
 “melons and gourds (as a class)”</p> <p>d. mjìtkhǎúm -
 untie
 kùmtìn
 cause to be untied
 “untie at once”</p> | <p>e. ñpuj - lǎzu
 wind gale
 “strong wind”</p> <p>(172) high < mid</p> <p>a. kjín - kǎtép
 hurry force
 “crisis”</p> <p>b. tìp - sep
 press exploit
 “exploit”</p> <p>c. kjìt - oí
 tie hang up
 “hang”</p> <p>d. mǎsù? - khǎlém
 deceive fool
 “hoodwink”</p> |
|--|--|

- e. phúj - pónép
wear place under
“bedding”
- (175) **mid** < **low**
- f. m̀̀s̀̀ù? - m̀̀k̀̀ò?
decive crooked
“false”
- a. l̀̀ko - l̀̀t̀̀á?
foot hand
“hands and feet”
- (173) **high** < **low**
- a. ṭji - khai
grandfather grandmother
“maternal grandparents”
- b. phún - kówá
tree bamboo
“bamboo and trees”
- c. thóm - pharj
behind after
“future”
- (176) **low** < **low**
- a. ph̀̀ùj - f̣iŋkàŋ
prestige dignity
“prestige”
- (174) **mid** < **mid**
- a. ǹ̀n - ts̀̀ò
low high
- a. f̀̀àn - ŋá
flesh fish
‘feel dizzy, confused’
- b. s̀̀ài - f̀̀àn
blood flesh
‘flesh and blood; kin’

5.5.1.2 Analysis

No matter what analysis of these data is accepted, it must be granted (extra-theoretically) that phonological factors can influence the surface linear order of morphological constituents. This influence could take one of at least four forms:

- (177) a. **Linearization mechanism.** The morphosyntax may pass incompletely linearized structures to the phonology, which is then responsible for determining their linear order.
- b. **Filter on derivation.** The morphosyntax may generate fully-linearized forms with both sequences; the phonology filters out those that do not ac-

cord with the ordering criteria.

- c. **Readjustment mechanism** The morphosyntax passes fully-linearized strings to the phonology, which the phonology is able to manipulate.
- d. **Deciding constraint in strongly parallel grammar.** Phonological constraints act in parallel with morphosyntactic constraints, either overriding these constraints or determining linear order when there is no overriding morphosyntactic constraint.

In this analysis, and the analyses that follow, we will assume (177d) but the analysis in most of its details will also work with all of the models except for (177b). In the case of (177b), the phonology must be able to generate absolute ungrammaticality/ineffability. While this is possible in Optimality Theory (Prince and Smolensky 1993; Orgun and Sprouse 1999), none of the mechanisms proposed to do so is fully adequate and in the spirit of the theory as a whole. In order to function as a filter, the phonology must contain some sort of “hard” constraint. For the purposes of our analysis, though, these issues are only of tangential importance.

What is important is the question of whether a scale is required to properly analyze this set of data. If we assume that phonological constraints can determine the sequence of morphosyntactic constituents, a first approximation of a model might rest on the fact that at most one of the conjuncts can be in final position. We could posit constraints that apply to the final vowel in a word: $*[+high]_{\sigma\#}$, penalizing final-syllable vowels with the feature $[+high]$, and $*[-low]_{\sigma\#}$, penalizing final-syllable vowels with the feature $[-low]$. If $*[+high]_{\sigma\#}$ dominates $*[-low]_{\sigma\#}$, then we predict the kind of ordering effect seen:

		/kàtéɸ, kjín/	$[+high]_{\sigma\#}$	$[-low]_{\sigma\#}$
(178)	a.	(a) kàtéɸ-kjín	*!	
		☞ (b) kjín-kàtéɸ		*

		/lòko, lətáʔ/	[+high] _{σ#}	[-low] _{σ#}
b.	☞ (a)	lòko-lətáʔ		
	(b)	lətáʔ-lòko		*!

At first, this may seem like an elegant solution. It only requires competition between two constraints, both of which make reference only to features that are widely accepted, [high] and [low]. It also has not required making reference to correspondence relationships. This analysis has some serious defects, however.

In the first case, this analysis requires positing of bizarrely ad hoc constraints, which make arbitrary reference to a structural position without any type of general theoretical motivation. Perhaps more seriously, this analysis divides the motivations for ordering into two unrelated constraints. Ordering effects could also be derived by the effect of constraints like *_{σ#}[+anterior], so that conjuncts with labial or coronal consonants in the final syllable tend to be ordered in first position. This state of affairs obscures Dai's clear insight regarding this phenomenon: that it is driven by the comparative height of the tonic vowels. In this formulation, it is a single, unified principle that motivates the sequencing pattern, and this principle calls for a unified explanation.

This is possible if we assume the existence of a scale, corresponding to the three levels of vowel height:

$$(179) \quad S = \{a\} > \{e, o\} > \{i, u\}$$

We have already argued for the existence of correspondence relationships across forms, and that these correspondence relationships are preferentially between constituents in prosodically prominent positions. In Chapters 3–4, we also argued for the existence of DIRECTIONAL ANTI-FAITHFULNESS, on the basis of a variety of chain-shifting phenomena. If we apply these two—independently motivated—concepts to the problem of compound ordering in Jingpho, we can analyze the relevant data with no additional theoretical stipulations. While such an analysis may seem somewhat more complicated,

it is actually simpler in the sense that it involves adding no novel machinery to the grammar.

As it is formulated in Chapter XYZ, WAX is the string-internal analogue of HIGHER. It may be defined as follows:

(180) *WAX[S]

Let α and β be corresponding entities within an output string. If α and β are in the domain of S and $\alpha \prec \beta$ then $\alpha < \beta$ in S .

In informal terms, this means that when two constituents (in the Jingpho case, two vowels) are in correspondence, then the second should be higher along the scale than the first. When applied to our case, this may seem wrong at first, due to the fact that high vowels are at the *bottom* of the Jingpho vowel scale in (179). When viewed at a sufficient level of abstraction, the only difference between this constraint and HIGHER is the fact that linear relationships are substituted for input-output relationships. Once it is understood that both types of correspondence have inherent in them an equivalent kind of asymmetry, as argued at length by (Hansson 2001), and what is treated as linear precedence in (180) is replaced with this asymmetry relation, the two constraints become formally equivalent. They will be referred to by different names only because their underlying identity is not easy to grasp at first and because individual constraints of this type are conceived of as applying to only one type of correspondence relationship.

Once we realize that this formal equipment is already available, the solution of the Jingpho problem is trivial. If $\text{CORR-}\sigma \leftrightarrow \sigma$ dominates WAX[S], where S is the scale in (179), the correct orderings are predicted:

		/kàté̃p, kjín/	CORR- $\acute{\sigma}$ \leftrightarrow $\acute{\sigma}$	WAX[S]
(181)	(a)	kàté̃p-kjín	*!	
	(b)	kàté̃ _i p-kjí̃ _n		*!
	(c)	kjín-kàté̃p	*!	
	(d)	kjí̃ _n -kà _i té̃ _i p		*!*
	(e)	☞ kjí̃ _n -kàté̃ _i p		
		/l̀̀ko, l̀̀tá̃ʔ/	CORR- $\acute{\sigma}$ \leftrightarrow $\acute{\sigma}$	WAX[S]
(182)	(a)	l̀̀ko-l̀̀tá̃ʔ	*!	
	(b)	☞ l̀̀ko _i -l̀̀tá̃ _i ʔ		
	(c)	l̀̀tá̃ʔ-l̀̀ko	*!	
	(d)	l̀̀tá̃ _j ʔ-l̀̀ko _j		*!
	(e)	l̀̀tá̃ʔ-l̀̀ko	*!	*
	(f)	l̀̀tá̃ _i ʔ-l̀̀ko _i		*!

If WAX was ordered above CORR- $\acute{\sigma}$ \leftrightarrow $\acute{\sigma}$ and the other CORR constraints, then—perhaps paradoxically—it would not typically have an effect upon ordering because correspondence relationships would only be established when they did not lead to WAX[S] being violated.

There is another possible analysis. De Lacy (2002a) has argued that there is a universal affinity that exists between sonorous segments and prosodically prominent positions. He encodes this relationship in a family of markedness constraints, which are in a stringency-relation with one another (relative to a scale). This means, in effect, that any constraint against low vowels in head positions (DTEs) must also penalize mid vowels and high vowels in those positions. Likewise, constraints against mid vowels in head positions must also penalize high vowels. There are symmetrical sets of constraints: those that penalize more sonorous vowels in non-head positions and those that penalize less sonorous vowels in head positions. These constraints use the following

notation:

- (183) a. $*\Delta_{PrWd} \leq \{i, u\}$
 “No high vowels in the heads of prosodic words.”
 b. $*\Delta_{PrWd} \leq \{i, u, e, o\}$
 “No high or mid vowels in the heads of prosodic words.”
- (184) a. $*-\Delta_{PrWd} \geq \{a\}$
 “No low vowels outside the heads of prosodic words.”
 b. $*-\Delta_{PrWd} \geq \{a, e, o\}$
 “No low or mid vowels outside the heads of prosodic words.”

If we assume that Jingpho prosodic words have a head at their right periphery (a proposition for which we have no direct evidence) it becomes immediately evident that we can generate ordering effects of the type we have observed here.

		$*\Delta_{PrWd} \leq \{i, u\}$	$*\Delta_{PrWd} \leq \{i, u, e, o\}$
(185) a.	(a) lətáʔ-ləkɔ		*!
	☞ (b) ləkɔ-lətáʔ		
		$*\Delta_{PrWd} \leq \{i, u\}$	$*\Delta_{PrWd} \leq \{i, u, e, o\}$
b.	(a) kətép-kjín	*	*!
	☞ (b) kjín-kətép		*

This analysis seems to account for the data in (171-176) just as well as the analysis stated in terms of correspondence constraints and distantal anti-identity. However, there are important reasons for accepting the earlier analysis rather than the analysis stated in de Lacion terms. First, it should be noted that the de Lacion analysis depends crucially upon the supposition that the head of the prosodic word in Jingpho is at the right periphery (and, indeed, the assumption that Jingpho co-compounds are coextensive with prosodic words). There is no independent evidence for this proposition except for the apparent fact that stress feet in Jingpho are iambic, and thus right-headed.

As additional examples of ordering effects are introduced, it will become clear that this analysis cannot be extended to most of the other cases we will examine. There are two reasons that this is the case. First, in contrast to the Jingpho case (and the Tangkhul case which we will discuss next), most of the cases we will examine cannot be explained in terms of a “universal” scale like sonority. Since the whole purpose of de Lacy’s framework is to constrain possible markedness relationships, and since it contains the explicit claim that markedness scales are irreversible, it cannot be applied profitably to these cases. Second, under this type of analysis, the relevant markedness constraints can only favor more sonorous vowels in the heads of feet and prosodic words. This cannot explain why, in some languages we will examine below, the first syllable of a conjunct is treated the same way as the stressed syllable, only more so. Thus, while an analysis stated in terms of constraints like de Lacy’s looks promising initially, it is ultimately insufficient to account for this phenomenon, and need not be invoked even for compatible cases, like that of Jingpho.

5.5.2 Comparison to Tangkhul

In Tangkhul, a Tibeto-Burman language of Manipur State, India, there appears to be an ordering effect that is remarkably similar to that seen in Jingpho. In all available examples of coordinate compounds, the last vowel in the first conjunct is higher than that in the second conjunct. The data are not adequate to make a strong claim, but to the extent that data are available, the generalization is perfect. Take the following data from Ahum (1997b,a):

(186) **high** < **high**

- a. kə-lí - kə-súí
 NOM-steal NOM-bait
 ‘stealing/cheating, etc.’

- b. kə-tʰi - kʰə-rɪŋ
 NOM-die NOM-live
 ‘by hook or by crook/by any means’

(187) **high** < **mid**

- a. zék-kè-fì -
face-NOM-bad
zèk-kè-kòr
face-NOM-ugly
'ugly/abominable'

(188) **high** < **low-mid**

- a. kè-pí - kè-pèm
NOM-sleep NOM-sit
'lodging, etc.'
- b. luŋ-cí - luŋ-hèr
stone-salt stone-ball
'any crystal stone'

(189) **high** < **low**

- a. k^hè-ŋè-cím - k^hè-ŋè-naj
NOM-hate NOM-slippery
'utter hate/dislike, etc.'
- b. cì-ftú - cìhá
dust scrap/dust
'dust/scrap'
- c. kè-t^hì - kè-sàr
NOM-die NOM-old
'death, calamity, etc.'
- d. k^hè-ŋè-mú - kè-ŋè-sáw
NOM-fight NOM-shout
'fighting/turmoil, etc.'
- e. kè-fì - kè-p^há
NOM-bad NOM-good
'anything (good or bad)'
- f. k^hè-uŋ - k^hè-và
NOM-come NOM-go
'transportation/journey, etc.'

- g. à-vú - à-và
PFX-mother PFX-father
'parents'

- h. ŋè-sún - ŋè-já
day night
'day and night/most of the
time'

(190) **mid** < **mid**

- a. kè-t^héj - kè-t^hem
NOM-know NOM-skill
'skillful/expert, etc.'
- b. k^hè-rèj - kè-tew
NOM-big NOM-small
'everybody'

(191) **mid** < **low-mid**

- a. k^hè-rèj - k^hè-ŋèm
NOM-huge NOM-stubborn
'powerful/big shot, etc.'
- b. haŋ-sók - haŋ-zəŋ
speak-out speak-in
'be frank/sincere/outspoken'

(192) **mid** < **low**

- a. k^hè-ŋ-p^héj - k^hè-ŋ-p^haŋ
plate-leg plate-branch
'utensils'
- b. kè-cot - kè-càŋ
NOM-tire NOM-tire
'weariness/suffering, etc.'
- c. à-wò - à-và
PFX-grandfather PFX-father
'forefathers'

(193) **low-mid** < **low-mid**

- a. k^hə-rér - nào-sən
NOM-old child-DIM
'everybody (young and old)'

(194) **low-mid** < **low**

- a. kə-zət - kə-kàr
NOM-go NOM-tread
'journey/way, etc.'
- b. kə-fək - kə-zà
NOM-drink NOM-eat
'foot/diet, etc.'
- c. kə-k^hə-ŋəŋ - kə-kə-zà
NOM-tire NOM-sick
'sickness, trouble, etc.'

(195) **low** < **low**

- a. kə-sá - kə-và
NOM-do NOM-go
'deed/movement'
- b. sɪna - lùpa
gold money
'money'
- c. kə-p^há - k^hə-jà
NOM-good NOM-good
'good (behavior/character)'
- d. mə-jar - ŋə-là
boy girl
'youth'

The data from Tangkhul is suggestive in its absolute regularity, but problematic in its scarcity¹⁵. In fact, if the vowel phonemes are simply plotted against one another, it is not entirely clear that there is a pattern at all:

¹⁵Indeed, I cannot make a firm statement about the relative frequency of these forms. While I spent over a year working with speakers of Tangkhul, I only came across a few forms of this type—four of which were actually not from Tangkhul-proper, but from closely-related languages—and would have been unaware that the construction was important in Tangkhul without the information provided by Ahum (1997b,a). If this construction is productive, it is, nevertheless, not easy to elicit.

(196) **Types of Tangkhul coordinate compounds by vowels**

V1\V2	i	u/ʊ	e	o	ɐ	a
i	1			1	2	3
u/ʊ						5
e			2		1	1
o					1	2
ɐ					1	2
a						4

However, if the vowels are reduced to four classes, defined in terms of vowel height, a pattern appears quite clearly:

(197) **Types of Tangkhul coordinate compounds by vowel height**

V1\V2	high	mid	lmid	low
high	1	1	2	8
mid		2	2	3
lmid			1	2
low				4

This fact suggests a scale with precisely this structure:

- (198) a. $i, u \prec e, o \prec \text{ɐ} \prec a$
 b. $\{i, u\} < \{e, o\} < \{\text{ɐ}\} < \{a\}$

It is also important to note that, even given this small number of data, the probability that this generalization is the result of chance—that it just happens to be the case that our data-set is biased—it should be noted that the chance of having eighteen tokens with the higher vowel first, where ordering is otherwise free, and getting no tokens where the higher vowel is last, is vanishingly small. Ordering a co-compound is no different, from the standpoint of probability, from tossing a coin; and detecting an ordering bias is the

same as distinguishing a fair coin from a biased coin (at some abstract level, at least). Since the probability of the higher vowel being ordered first, under this view of things, is $1/2$, the probability of finding eighteen tokens of this type, and no tokens of the other, is $1/2^{18}$ which is approximately 3.8×10^{-6} or one in 262,144¹⁶. The probability of high vowels being ordered before lower vowels (mid, low-mid, and low) is also exceedingly slim, begin approximately 4.9×10^{-4} , and the probability of mid vowels appearing before low vowels five times with no instances of the reverse type is also quite small: 0.03 or one in 32. It is hard to dispute that this generalization is real, and based upon the available data, there is no reason to believe that it is not the same type of hard ordering effect seen in Jingpho.

The formal analysis is exactly the same as that in Jingpho—not surprising, given the structural similarity of the two languages. In fact, it should not escape our notice that we have now seen three different cases of vowel-driven co-compound ordering constraints (or tendencies) in three different languages belonging to the Tibeto-Burman family, each of which belongs to a separate branch of that family. It has been suggested (e.g. Matisoff 1991) that Jingpho and the Lolo-Burmese family (to which Lahu belongs) may belong to the same top-level subgroup in the Tibeto-Burman family. However, if such a relationship does exist, it is a very distant one. Tangkhul, by all accounts, belongs to a completely different subgroup in Tibeto-Burman (which the author believes to include the Kuki-Chin languages as well, based upon shared sound-changes such as PTB *s- →t and PTB *ts > s). This suggests that either the same pattern has developed independently in these three languages, or that a pattern of this sort already existed in Proto-Tibeto-Burman and comes to Jingpho and Tangkhul as common inheritance shared by the family as a whole. Far more substantive work would be required in

¹⁶There are actually more sophisticated ways of reasoning about the probabilities of events of this type. This simple model is employed here because it is easy to explain and makes the point effectively, and not because the other methods would produce contradictory results.

order to determine if this was truly the case, but if it is, we must view the vowel-based ordering tendencies in Lahu not to be an inchoate system like Lahu's tone-driven ordering preferences, but the waning remnants of a once-binding law.

5.6 Co-Compounds in Hmong (Mong Leng)

We have already seen that hard ordering effects can be driven by vowel quality (Jingpho and possibly Tangkhul) and that soft ordering effects can be driven by both vowel quality (Lahu) and tone (Chinese and Lahu). We will now see a case of a tone ordering generalization that is different from that in Lahu in a few respects. While the Lahu effect is apparently quite new, the effect we will now discuss had to have originated in the distant past; while the Lahu effect was phonetically transparent, a monotonic function over tone height, this pattern is largely (though not entirely) arbitrary, from a phonetic standpoint; and, while the influence of tone upon the organization of co-compounds is—as of now—purely a preference or tendency, this effect—in Hmong—has the status of a (productive) grammatical rule.

In all the dialects of Hmong¹⁷ for which documentation is available, there is an ordering effect in coordinate compounds. Like the second type of effect found in Lahu and the effect noted in Chinese, this generalization is based upon tonal contrasts. Like the Jingpho case, however, co-compound ordering in Hmong is categorical.

¹⁷As the term *Hmong* is used here, it refers to a specific group of dialects in the Far Western Hmongic group (the 川黔滇 Chuanqiandian group) that are mutually intelligible: Hmong Daw (White Hmong), Mong Leng (Green/Blue Hmong), and the dialects of 大南山 Dananshan and Xuyong in China. This would exclude the more closely-related languages of the A-Hmao (滇东北 Diandongbei) group as well as the more distantly related 麻山 *Mashan* group and the even more distantly related languages/dialects of the 羅泊河 *Luopohe*, 重安江 *Chong'anjiang*, 貴陽 *Guiyang*, and 惠水 *Rongshui* groups, all of which are sometimes referred to as *Hmong*. In fact, the whole Hmongic branch (with the exception of Bunu, Baheng, and so forth) is sometimes referred to by this label, but a much more restricted use of the term is adopted here.

Here we will look at Hmong co-compound ordering using data from Mong Leng, a dialect of Hmong spoken in parts of China, Vietnam, Laos, and China. As was discussed in Chapter 3, Proto-Far Western Hmong had eight contrasting tones: four historical tones divided into two registers.

(199) Far-Western Hmongic tonal categories

Register\Tone	A	B	C	D
1 (“high”)	A1	B1	C1	D1
2 (“low”)	A2	B2	C2	D2

In some Hmong dialects, e.g. Dananshan¹⁸ (Wang 1985, 1979, 1994) these eight tonal categories are retained intact.

(200) Dananshan tones

Register\Tone	A	B	C	D
1 (“high”)	HM	H↑	H	M
2 (“low”)	ML	L (breathy)	LM	MH
			(breathy)	

In Mong Leng, however, two of the categories have merged:

(201) Mong Leng tones

Register\Tone	A	B	C	D
1 (“high”)	H (x̂)	MH (x̃)	M (x)	L (ẋ)
2 (“low”)	HL (x̂)	ML (x̃)	ML (x̃)	L (ẋ)

Despite this fact, for purposes of exposition, these two categories are differentiated here, in the same way they are distinct in the Dananshan dialect. Some implications of

¹⁸Technically, Dananshan is a “Mong Leng” dialect. However, it differs in a number of respects from the dialect to which that name has typically been applied.

this merger for the behavior of modern speakers of Mong Leng are discussed below.

5.6.1 The ordering generalization

Like the related tendency in Chinese, the ordering generalization in Hmong is not as easy to express as those in Lahu and Tangkhul. Because of the fact that its organizational principles are not directly tied to phonetic substance, it is not possible to reduce generalization to a statement about pitch height or any other phonetic dimension. It is probably best illustrated visually, by viewing co-occurrence distributions. In the following table, the tonal combinations are given for a corpus of 198 co-compounds¹⁹. The tones of the first conjunct are classified by row and the tones of the second conjunct are classified by column:

¹⁹The collection of compounds was derived from a number of sources: my own fieldwork, particularly in the years from 1995–1997, published texts, informal texts obtained from Internet fora, and lexicographical collections (Xiong et al. 1992; Lyman 1974). The forms and glosses have been checked against other references including Heimbach (1979); Lyman (1974); Xiong et al. (1992); Xiong (2005).

(202) Occurrence of tonal pairs in Mong Leng coordinate compounds

	\hat{x}	\acute{x}	\underline{x}	\grave{x}	$\underline{\underline{x}}_1$	\check{x}	$\underline{\underline{x}}_2$	x
\hat{x}	1	5	6	3	3	16	22	10
\acute{x}		6	6	8	2	13	12	24
\underline{x}			2	1	—	5	4	5
\grave{x}				2	1	2	—	4
$\underline{\underline{x}}_1$				1	1	1	3	4
\check{x}						2	4	9
$\underline{\underline{x}}_2$			1	1	1	1	3	—
x								3

Of the 198 compounds in the collection, 97.5% follow the generalization implied by the table. That is to say, they are ordered such the first conjunct is never higher than the second conjunct along the following scale:

(203) Observable tone scale in Mong Leng

- a. $\hat{x} \downarrow < \acute{x} \uparrow < \underline{x} \downarrow < \grave{x} \downarrow < \underline{\underline{x}}_1 \downarrow < \check{x} \uparrow < \underline{\underline{x}}_2 \downarrow < x \uparrow$
- b. $HL < H < \underline{L} < L < \underline{\underline{L}} < MH < \underline{L} < M$

Data illustrating these distributions (as exhaustively as possible) are given below:

(204) FALLING TONE FIRST

a. **falling** \prec **falling**

- i. $\int\hat{i}$ -tsî -
RECIP-as tall
 $\int\hat{i}$ -lûa
RECIP-as much
'equal'

b. **falling** \prec **high**

- i. ntû - té
heaven earth
'heaven and earth; the universe'
- ii. lâ - té
rice paddy dry field
'agricultural land; agriculture'

c. **falling** < **creaky**

- i. ɲû - t̥i
cow buffalo
'bovines; large livestock'
- ii. tsî - nkai
match pair
'spouse'

d. **falling** < **low**

- i. nô - hau
eat drink
'consume; sustenance'
- ii. jâŋ - tʃ^hi
sheep goats
'ovines'

e. **falling** < **breathy**₁

- i. tâ - t̥u
level settled
'even; calm; peaceful'
- ii. ɲû - neŋ
cow horse
'livestock'

f. **falling** < **rising**

- i. tlê - cǎi
water whisky
'liquor; drink'
- ii. t̥ô nts^hǎŋ
oil blood
'blood; bodily liquids'

g. **falling** < **breathy**₂

- i. nkâu -
male youth
n̥tau
female youth
'young men and women;
romance'
- ii. n̥tô - t̥ɟua
herbs bushes
'low-lying vegetation'

h. **falling** < **mid**

- i. nâ - ɕoŋ
year year
'year'
- ii. nqâ - no
meat rice
'food; feast'

(205) HIGH TONE FIRST

a. **high** < **high**

- i. fúa - tséŋ
plunder seize
'pillage'
- ii. tú - n̥aiŋ
son daughter-in-law
'sons and daughters-in-
law'
- iii. hlú - páŋ
love help
'come to the aid of' (also
páŋ-hlú)

b. **high** < **creaky**

i. lá - cua
monkey gibbon
'simians'

ii. káj - ntʃau
bug ant
'insects'

c. **high** < **low**

i. tláj - q^húa
spirit guest
'ghosts and guests (do-
mestic spirits)'

ii. q^húa - ntúa
instruct exhort
'counsel'

d. **high** < **breathy**₁

i. fá - qe
high low
'high and low; every-
where'

ii. jóŋ -
high-pitched
ləu
low-pitched
'high-pitched and low-
pitched'

e. **high** < **rising**

i. káj - kě
path way
'ceremony'

ii. kóŋ - mǝŋ
fortune fate
'fortune'

f. **high** < **breathy**₂

i. néŋ - jǝi
shamanism divination
'shamanistic practices'

ii. mó - nkeŋ
hurt tired
'ill; illness'

g. **high** < **mid**

i. qá - npua
chicken pig
'small livestock'

ii. hlú - tʃ^hua
love care for
'cherish'

iii. cóŋ - nton
bamboo tree
'large, woody vegetation'

(206) CREAKY TONE FIRST

a. **creaky** < **creaky**

i. ts^hu - ts_i
fault guilt
'blame; wrongdoing'

ii. xǝŋ - pu
evaluate see
'appraise'

b. **creaky** < **low**

- i. p^hẹ - qà
bad disgusting
'vile'

c. **creaky** < **breathy**₁

- i. none

d. **creaky** < **rising**

- i. jaŋ - tǣŋ
kind sort
'kind'
- ii. nu - tǣ
official lord
'rulers; leaders'

e. **creaky** < **breathy**₂

- i. t^h aŋ - nu
talk ask
'be in touch'
- ii. tʃh_um - nkau
poke stab
'stab'

f. **creaky** < **mid**

- i. tǣ - tʃaŋ
create raise
'build up; develop'
- ii. ts^hị - k^hu
mend fix
'repair; restore'

(207) **LOW TONE FIRST**

a. **low** < **low**

- i. tau - tsua
axe bush knife
'chopping tools'
- ii. tsua - tlhò
join fit together
'assemble; put together'

b. **low** < **breathy**₁

- i. hlua - lau
young old
'young and old; everyone'

c. **low** < **rising**

- i. ʃitxi -
RECIP-marry
ʃiǰüa
RECIP-marry
'be married to one another'

d. **low** < **breathy**₂

- i. none

e. **low** < **mid**

- i. ti - tai
wing foot
'wings and feet; appendages of bird'
- ii. ntàu - tua
hit kill
'fight; do battle'

(208) BREATHY₁ TONE FIRST

a. **breathy₁** < **breathy₁**

- i. moŋ - lu
go come
'come and go; back and forth'

b. **breathy₁** < **rising**

- i. tai - tlã
dish spoon
'dishes; tableware'

c. **breathy₁** < **breathy₂**

- i. naŋ - noŋ
rodent bird
'small animals'
- ii. ntsa - tlaŋ
entice lie
'deceive'

d. **breathy₁** < **mid**

- i. te - tai
hands feet
'hands and feet; appendages; martial arts'
- ii. tɕi - tɕho
pants shirt
'clothing'

(209) RISING TONE FIRST

a. **rising** < **rising**

- i. chĩ - kũ
touch touch
'touch'

- ii. tĩ - t^hãi
resist protect
'defend'

b. **rising** < **breathy₂**

- i. fã - ʒu
life strength
'effort'

- ii. tɕũ - tɕo
tiger war
'war'

c. **rising** < **mid**

- i. cãŋ - tɕhe
contend curse
'castigate'
- ii. t^hõ - cai
entreat invite
'invite (e.g. to act in some official capacity)'

(210) BREATHY₂ TONE FIRST

a. **breathy₂** < **breathy₂**

- i. tlaŋ - ʒu
strength strength
'strength'

- ii. taŋ - kai
end end
'end (v.)'

b. **breathy₂** < **mid**

- i. none

(211) MID TONE FIRST

a. **mid** < **mid**

i. k^hu - q^ha
fix teach

‘correct (someone’s) be-

havior’

ii. ts^hua - ts^ha
every every
‘every’

With very few exceptions, it is possible to predict the order of the ordinary type of coordinate compounds based wholly upon the tones of the two conjuncts. There are two classes of exceptions to this principle, the first (and most obvious) one being the set of co-compounds where the tones of the conjuncts do not differ. In these cases, some other principle must make the final determination of sequence. In some of these cases, the order is, in fact, indeterminate. Thus, both *hlú-páj* ‘love-help’ and *páj-hlúb* ‘help-love’ are licit co-compounds. In most cases of “well-established” co-compounds, there is a conventionalized order that seems to be part of the lexical representation of the compound as a word. There seem to be a number of influences on the sequence of conjuncts in situations of this type, the most obvious of which is the tendency, particularly in synonym-synonym compounds, for a rare or bound root to occur before a common one. Thus, since *fúa* ‘plunder’ is a relatively infrequent word and *tséŋ* ‘seize’ is very common, we correctly predict that they should be conjoined as *fúatséŋ*. Likewise, if there is a “natural” sequence of elements to which the two conjuncts belong, like the series of numerals, then the sequence of conjuncts will follow this ordering. Thus, there are co-compounds *ó-pé* ‘two-three’ and *tau-çaj* ‘six-seven’, but no compounds *pé-ó* or *çaj-tau*. However, this constraint is less “powerful” than the tonal ordering constraint so that the sequence of *jí* ‘eight’ and *cûa* ‘nine’ is *cûa-jí* ‘nine-eight’ rather than *jí-cûa* ‘eight-nine’. The ordering generalizations behave as if there are some semantic constraints which are, to borrow a metaphor from Optimality Theory, dominated by the tonal ordering constraints.

The tonal ordering generalization itself turns out to be slightly more complicated than the generalization that has been expressed thus far. It should have been noticed that some of the coordinate compounds in (204–209) have conjuncts with two syllables, typically the reciprocal prefix *fī-* added to some verb. In these cases—and in elaborate expressions of the ABAC type—the tones of the final syllables of each conjunct determine the organization of the construction in the same way that they would if they were the only tone in each conjunct. Thus, there are constructions like *kāi-ntáu-kāi-ntāi* ‘be educated (“study-cloth-study-paper”)’ but no constructions like **kāi-ntāi-kāi-ntáu* ‘study-paper-study-cloth’. If a construction is of the pseudo-elaborate expression type, with an ABCD structure, then both the (A, C) and (B, D) pairs will be ordered as if they are co-compounds by themselves, as will be discussed in greater detail below. If co-compounds have two (or more) syllables in each conjunct but these are not in the “free-assortment” type of relationship associated with pseudo-elaborate expressions, and if the tones of the first syllables of the two conjuncts differ, this difference will have precedence in determining the relative order of the conjuncts, to a first approximation (it will be shown below that the reality is slightly more complicated than this). Thus we have *pūjai-tsǐtsóŋ* ‘grandparents-ancestors’ rather than **tsǐtsóŋ-pūjai* ‘ancestors-grandparents’.

Despite these complications, it is evident that the ordering generalization that has been described here, or a generalization so close to it as to be formally indistinguishable, is both robust and correct. On the other hand, this is troubling since the scale that we derived from this generalization appears to be arbitrary. The progress through the categories *falling, high, low creaky, low, low breathy₁, rising, low breathy₂, mid* seems like nothing more than Brownian motion about the phonetic ocean. It appears, at first blush, that the organization of this scale and the implicit relationships it contains are purely random.

5.6.2 Factoring the observed scale

At a deeper level, though, the scale we have arrived at through superficial analysis of the data is anything but arbitrary. The structure of the scale is obscured, though, by the fact that we have concentrated on the phonetic exponence of the tonal categories involved, and not how they relate to one another. If we replace these substantive labels with the historical labels that were introduced at the beginning of this section, a clear pattern emerges:

$$(212) \quad A2 < A1 < D2 < D1 < B2 < B1 < C2 < C1$$

This scale can be divided into four pairs of contiguous categories, each of which corresponds to one of the four historical tones:

$$(213) \quad \{A2 < A1\} < \{D2 < D1\} < \{B2 < B1\} < \{C2 < C1\}$$

Within each of these pairs, the category belonging to the “low” register always precedes that belonging to the “high” register. This is significant not only because it begins to reveal the diachronic order that underlies the apparent chaos of the scale we have distilled from the data, but also because it interacts with an independently motivated part of the phonological grammar.

5.6.2.1 Independent evidence for the “register” scale

In Hmong tonology, it is necessary to recognize the synchronic existence of the diachronic distinction we have called register. By way of review, the register distinction was originally a voicing distinction in onsets. Syllables with voiceless and preglottalized onsets patterned together as the “high” register; other syllables (those with plain voiced onsets) formed the “low” register. This distinction has the following synchronic reflexes:

- (214) a. **Tone sandhi** There are two different tone sandhi patterns in Hmong, and these correlate with the register distinction. In the “high” register, there is chain lowering. In the low register, there is neutralization to the low breathy tone. Tone sandhi processes *never* result in register changes in Mong Leng.
- b. **Distributional restrictions** In native vocabulary (excluding expressives) aspirated onsets (and voiceless sonorants, in dialects that preserved them) do not appear in syllables in the “low” register.
- c. **Phonetic restrictions** Register distinctions line up with marked laryngeal features. Marked laryngeal features (non-modal voicing) never occur with tones in the “high” register. All tones in the low register show a marked laryngeal setting (tense phonation for the *falling* tone, and for the *breathy* and *creaky* tones, breathy and creaky phonation)

Because of the way these generalizations align, it is apparent that a grammar that does not capture these relationships is missing a significant generalization. In fact, for a variety of reasons, it is not possible to construct a viable model of the tone sandhi processes in Hmong without access to this distinction.

If register is treated as a scale, we can factor the superficial scale we constructed above into two simpler scales, *R* for “register” and *T* for “tone”:

- (215) a. $R = \{\hat{x}, \underline{x}_1, \underline{x}_2, \underline{x}\}_0 < \{\acute{x}, \grave{x}, \check{x}, x\}_1$
- b. $T = \{\hat{x}, \acute{x}\}_0 < \{\grave{x}, \underline{x}\}_1 < \{\underline{x}_1, \check{x}\}_2 < \{\underline{x}_2, x\}_3$

Given these two much simpler scales, and WAX constraints referring to them, explaining the basic ordering pattern is quite simple.

5.6.2.2 A competition between two WAX constraints

The two constraints which must be posited are WAX[*R*] and WAX[*T*]. The basic insight is that a co-compound is ordered according to *T* unless this is impossible (because the

two tones are not distinguished by T); under this condition, the relative ordering is determined by R . This indicates that $WAX[T]$ must dominate $WAX[R]$. These constraints, in turn, must be dominated by some constraint that enforces correspondence. Since, like Jingpho, Hmong has (according to the available evidence) iambic stress and right-headed prosodic constituents, it is fair to begin with $CORR-\acute{\sigma} \leftrightarrow \acute{\sigma}$, though it will soon become apparent that this is not adequate.

If we start with the ranking $WAX[R] \gg WAX[T] \gg CORR-\acute{\sigma} \leftrightarrow \acute{\sigma}$, we make the right predictions for compounds where each conjunct has a single syllable. The example for $ny-t\check{i}$ ‘leaders (“official-lord”’) demonstrates the outcome if the tones of the two conjuncts differ along the scale T :

		$CORR-\acute{\sigma} \leftrightarrow \acute{\sigma}$	$WAX[T]$	$WAX[R]$
(216)	(a)	$t\check{u}-ny$	*!	
	(b)	$t\check{u}_i-ny_i$		*!
	(c)	$ny-t\check{u}$	*!	
	(d)	$ny_i-t\check{u}_i$		*

Likewise, this ranking predicts the correct ordering for the case of a disyllabic compound where the two conjuncts do not differ relative to T but do differ relative to R :

		$CORR-\acute{\sigma} \leftrightarrow \acute{\sigma}$	$WAX[T]$	$WAX[R]$
(217)	(a)	$nt\hat{u}-t\acute{e}$	*!	
	(b)	$nt\hat{u}_i-t\acute{e}_i$		*
	(c)	$t\acute{e}-nt\hat{u}$	*!	
	(d)	$t\acute{e}_i-nt\hat{u}_i$		*!

However, the set of outputs that the grammar must generate is by no means limited to examples of this type. In §5.1, “pseudo-elaborate expressions” were mentioned. These compounds consist of four morphemes A, B, C, D, where (A, B) and (C, D) are subordinating grammatical structures (often subordinating compounds), which function as

unified prosodic domains relative to processes like stress assignment and tone sandhi. The pairs (A, C) and (B, D), at the same time, are in coordinate relationships. This means that there are always four possible arrangements of elements in such a compound (the product of the permutations of (A, C) and those of (B, D)). If we examine the case of *káj-tf^hój-kě-ku* ‘wedding rituals (“path-wedding-way-marriage”)', we can see that the current ranking cannot properly constrain this range of possibilities:

		CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[T]	WAX[R]
	(a) <i>kě-ku káj-tf^hój</i>	*!		
	(b) <i>kě-ku_i káj-tf^hój_i</i>		*!	
	(c) <i>kě_i-ku_j káj_i-tf^hój_j</i>		*!*	**
	(d) <i>kě-tf^hój káj-ku</i>	*!		
	✂ (e) <i>kě-tf^hój_i káj-ku_i</i>			*
	(f) <i>kě_i-tf^hój_j káj_i-ku_j</i>		*!	**
(218)	(g) <i>káj-ku kě-tf^hój</i>	*!		
	(h) <i>káj_i-ku kě_i-tf^hój</i>	*!		*
	(i) <i>káj-ku_i kě-tf^hój_i</i>		*!	
	(j) <i>káj_i-ku_j kě_i-tf^hój_j</i>		*!	*
	(k) <i>káj-tf^hój kě-ku</i>	*!		
	(l) <i>káj_i-tf^hój kě_i-ku</i>	*!		*
	✂ (m) <i>káj-tf^hój_i kě-ku_i</i>			*
	(n) <i>káj_i-tf^hój_j kě_i-ku_j</i>			**!

This ranking makes the right prediction for the (B, D) pair—the final syllables in the respective conjuncts—but underdetermines the sequence of the (A, C) pair. This is caused, quite obviously, by the fact that there is a constraint in the visible ranking that motivates a correspondence relationship between the heads of prosodic words but no constraint that encourages correspondence between the *initial* syllables of prosodic

words.

However, the need for such a constraint, $\text{CORR-}\#\sigma \leftrightarrow \#\sigma$, has already been discussed. If it is introduced into the ranking, above both $\text{WAX}[T]$ and $\text{WAX}[R]$, the resulting grammar rightly predicts the existence of $káj\text{-}t^h\acute{o}\eta\text{-}k\check{e}\text{-}k\grave{u}$ and the non-existence of $*k\check{e}\text{-}t^h\acute{o}\eta\text{-}káj\text{-}k\grave{u}$, since, other things being equal, it requires a correspondence relationship across both the (A, C) pair and the (B, D) pair:

		$\text{CORR-}\#\sigma \leftrightarrow \#\sigma$	$\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$	$\text{WAX}[T]$	$\text{WAX}[R]$
(a)	$k\check{e}\text{-}k\grave{u}\ káj\text{-}t^h\acute{o}\eta$	*!	*		
(b)	$k\check{e}\text{-}k\grave{u}_i\ káj\text{-}t^h\acute{o}\eta_i$	*!		*	
(c)	$k\check{e}_i\text{-}k\grave{u}_j\ káj\text{-}t^h\acute{o}\eta_j$			*!*	**
(d)	$k\check{e}\text{-}t^h\acute{o}\eta\ káj\text{-}k\grave{u}$	*!	*		
(e)	$k\check{e}\text{-}t^h\acute{o}\eta_i\ káj\text{-}k\grave{u}_i$	*!			*
(f)	$k\check{e}_i\text{-}t^h\acute{o}\eta_j\ káj\text{-}k\grave{u}_j$			*!	**
(g)	$káj\text{-}k\grave{u}\ k\check{e}\text{-}t^h\acute{o}\eta$	*!	*		
(h)	$káj\text{-}k\grave{u}_i\ k\check{e}_i\text{-}t^h\acute{o}\eta$		*!		*
(i)	$káj\text{-}k\grave{u}_j\ k\check{e}_i\text{-}t^h\acute{o}\eta_i$	*!		*	
(j)	$káj\text{-}k\grave{u}_j\ k\check{e}_j\text{-}t^h\acute{o}\eta_j$			*!	*
(k)	$káj\text{-}t^h\acute{o}\eta\ k\check{e}\text{-}k\grave{u}$	*!	*		
(l)	$káj\text{-}t^h\acute{o}\eta_i\ k\check{e}_i\text{-}k\grave{u}$		*!		*
(m)	$káj\text{-}t^h\acute{o}\eta_j\ k\check{e}_i\text{-}k\grave{u}_i$	*!			*
(n)	$káj\text{-}t^h\acute{o}\eta_j\ k\check{e}_j\text{-}k\grave{u}_j$				**

This is a promising result.

However, the situation is still more complex than the data and analysis presented thus far would suggest. There are also four-syllable coordinate compounds where one or both of the conjuncts is monomorphemic; in other similar cases, there is a special semantic association between morphemes within conjuncts such that the content of each conjunct is fixed and only the relative ordering of the conjuncts is free. In cases of this type, compounds are ordered—to a first approximation—according to the tones of the first syllable in each conjunct. Thus, $p\grave{u}j\grave{a}i$ ‘grandparents (“grandmother-grandfather”) and $ts\check{i}t\acute{s}\acute{o}\eta$ ‘ancestors’ (< Chinese 祖宗 $z\check{u}z\acute{o}ng$ ‘ancestors’) are compounded to form $p\grave{u}j\grave{a}i\text{-}ts\check{i}t\acute{s}\acute{o}\eta$ ‘ancestors’ rather than $ts\check{i}t\acute{s}\acute{o}\eta\text{-}p\grave{u}j\grave{a}i$, for although the high tone is lower

on *T* than the creaky tone, the falling tone is lower along *T* than the rising tone. The tones of the final syllable come into play only as a kind of “fall-back” mechanism when the tones of the initial syllables are identical. Thus, *pêseŋ* ‘citizenry’ (< Chinese 百姓 *bǎixìng* ‘hundred clans’) and *fúafǐ* ‘populace’ are compounded to form *pêseŋ-fúafǐ* ‘citizenry; subjects’ rather than **fúafǐ-pêseŋ*, since the falling tone is lower along *T* than the creaky tone. As long as $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ is ranked above $\text{WAX}[T]$, the ordering of these compounds will be incorrectly predicted (or more accurately, underdetermined), since $\text{WAX}[T]$ will not be able to motivate the violation of $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$. Take the case of *pûjai-tsítsóŋ*:

		$\text{CORR-}\#\sigma \leftrightarrow \#\sigma$	$\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$	$\text{WAX}[T]$	$\text{WAX}[R]$
	(a) tsũtsóŋ-pûjau	*!	*		
	(b) tsũtsóŋ _i -pûjau _i	*!			*
	☞ (c) tsũ _i tsóŋ _j -pû _i jau _j			*	**
(220)	(d) tsũ _i tsóŋ _j -pû _i jau		*!	*	*
	(e) pûjau-tsũtsóŋ	*!	*		
	(f) pûjau _i -tsũtsóŋ _i	*!		*	*
	(g) pû _i jau-tsũ _i tsóŋ		*!		*
	☞ (h) pû _i jau _j -tsũ _i tsóŋ _j			*	**

By reranking the constraints so that $\text{WAX}[T] \gg \text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$, we reach a grammar that makes the correct prediction about this form. This means that the initial syllables of each conjunct will always be in correspondence, regardless of whether this necessarily results in a violation of $\text{WAX}[T]$. However, $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ may be violated in order to reduce the violation profile of $\text{WAX}[T]$, capturing the desired asymmetry between the influence of tones in the two positions.

		CORR-# $\sigma \leftrightarrow \#\sigma$	WAX[T]	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[R]
(221)	(a)	tsũtsónj-pũjau	*!	*	
	(b)	tsũtsónj _i -pũjau _i	*!		*
	(c)	tsũ _i tsónj _j -pũ _i jau _j	*!		**
	(d)	tsũ _i tsónj _j -pũ _i jau	*!	*	*
	(e)	pũjau-tsũtsónj	*!	*	
	(f)	pũjau _i -tsũtsónj _i	*!	*	*
	(g)	pũ _i jau-tsũ _i tsónj		*	*
	(h)	pũ _i jau _j -tsũ _i tsónj _j	*!		**

Likewise, this ranking gives the correct outcome for *pêseŋ-fúafĩ* (noting that $\hat{x} = \acute{x}$ in *T* but $\hat{x} < \acute{x}$ in *R*):

		CORR-# $\sigma \leftrightarrow \#\sigma$	WAX[T]	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[R]
(222)	(a)	fúafũ-pêseŋ	*	*	*!
	(b)	fúafũ _i -pêseŋ _i	*!	*	*
	(c)	fúafũ _j -pêseŋ _j	*!*		**
	(d)	pêseŋ-fúafũ	*	*	
	(e)	pêseŋ _i -fúafũ _i	*!	*	*
	(f)	pêseŋ _j -fúafũ _j	*!*		*

The same facts are illustrated by *lâ?à-téc^hài* ‘country (“paddy-land”)’ where *lâ?à* is literally ‘paddy-paddy’ and *téc^hài* is literally ‘land-place’.

		CORR-# $\sigma \leftrightarrow \#\sigma$	WAX[T]	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[R]
(223)	(a)	lâ _i ?àtéc ^h àu	*	*	
	(b)	lâ _i ?à _i téc ^h àu _i	*!	*	*
	(c)	lâ _i ?à _j téc ^h àu _j	**!		*
	(d)	téc ^h àulâ _i ?à	*	*	*!
	(e)	téc ^h àu _i lâ _i ?à _i	*!	*	*
	(f)	téc ^h àu _j lâ _i ?à _j	*!*		**

It is best not to establish a correspondence relationship between the prosodic heads because this will satisfy CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ at the expense of violating the higher-ranked WAX[T] (as well as WAX[R]). This leaves the weight of the ordering decision upon

the tones of the stem-initial syllables. By ordering the compound so that *pêseŋ* is first and coindexing the stem-initial syllables, it is possible to avoid violating both CORR- $\#\sigma \leftrightarrow \#\sigma$ and WAX[R], while incurring only one violation each of CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ and WAX[T].

This new ranking still gives the right predictions for the cases we have examined so far, as demonstrated by the following derivations of *nutfĩ* ‘leaders (“official-lord”)’ and *káŋtʰóŋ-kěkũ* ‘marriage rites (“path-marriage-way-wedding”)’ (Note that, $\underline{x} < \check{x}$, $\acute{x} < \check{x}$, $\acute{x} < \underline{x}$ in *S*; $\underline{x} < \check{x}$, $\acute{x} = \check{x}$, $\underline{x} < \acute{x}$ in *R*).

		CORR- $\#\sigma \leftrightarrow \#\sigma$	WAX[T]	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[R]
(224)	(a) tʃũ-nũ	*!		*	
	(b) tʃũ _i -nũ _i		*!		
	(c) nũ-tʃũ	*!		*	
	☛ (d) nũ _i -tʃũ _i				*

		CORR- $\#\sigma \leftrightarrow \#\sigma$	WAX[T]	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[R]
(225)	(a) kê-kũ káŋ-tʃhóŋ	*!		*	
	(b) kê-kũ _i káŋ-tʃhóŋ _i	*!	*		
	(c) kê _i -kũ _j káŋ _j -tʃhóŋ _j		*!*		**
	(d) kê-tʃhóŋ káŋ-kũ	*!		*	
	(e) kê-tʃhóŋ _i káŋ-kũ _i	*!			*
	(f) kê _i -tʃhóŋ _j káŋ _j -kũ _j		*!		**
	(g) káŋ-kũ kê-tʃhóŋ	*!		*	
	(h) káŋ _i -kũ _i kê _i -tʃhóŋ			*!	*
	(i) káŋ-kũ _i kê _i -tʃhóŋ _i	*!	*		
	(j) káŋ _j -kũ _j kê _i -tʃhóŋ _j		*!		*
	(k) káŋ-tʃhóŋ kê-kũ	*!		*	
	(l) káŋ _j -tʃhóŋ _j kê _i -kũ _i			*!	*
	(m) káŋ-tʃhóŋ _i kê-kũ _i	*!			*
	☛ (n) káŋ _j -tʃhóŋ _j kê _i -kũ _j				**

This is a comforting confirmation that the analysis is on the right track.

More significantly, though, this grammar makes predictions regarding cases that have not yet been examined which turn out to be correct. One of these is that a tonal difference between the stem-initial syllables along *T* will be decisive in determining

the order of constituents in a four-syllable co-compound (where the structure of the conjuncts is fixed), but a difference along *R* will be less important than a difference along *T* in the prosodic head syllables. This is illustrated by the case where *vâŋtsǐ* ‘king’ (< Chinese 王子 ‘prince’) is compounded with *fúatài* ‘emperor’ (early loan from Chinese 黃帝 *huángdì*) to form *fúatài-vâŋtsǐ*. If both *R* and *T* were equally weighted, we would predict **vâŋtsǐ-fúatài* as the result, since the initial syllables have the tones \hat{x} and \acute{x} respectively, and $\hat{x} < \acute{x}$ relative to *R*. The grammar we have developed thus far makes the opposite (and correct) prediction:

		CORR-# $\sigma \leftrightarrow \#\sigma$	WAX[<i>T</i>]	CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[<i>R</i>]
(226)	(a) <i>vâŋ_intsǔ-fú_itài</i>		*	*!	
	(b) <i>vâŋtsǔ_i-fúatài_i</i>	*!	*		*
	(c) <i>vâŋ_jntsǔ_j-fú_itài_j</i>		*!*		*
	(d) <i>fú_itài-vâŋ_intsǔ</i>		*!	*	*
	(e) <i>fúatài_i-vâŋtsǔ_i</i>	*!			*
	(f) <i>fú_itài_j-vâŋ_jntsǔ_j</i>		*		**

Since it is more important to satisfy CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$ than to satisfy WAX[*R*], then candidate (a), where the tones of the stem-initial syllables would have determined the order of the conjuncts, must lose to (f), where the prosodic head syllables are co-indexed, allowing the difference along *T* between the tones of these syllables to determine the order of the conjuncts.

If we take these findings regarding coordinate compounds consisting of two “fixed” disyllabic conjuncts, and apply it to pseudo-elaborate expressions, we make another interesting prediction: if a compounding context triggers a tone sandhi process, this could potentially affect the order of conjuncts. We have already looked at the case of *ntû-té* ‘heaven and earth’ as a simple co-compound. Independently, *ntû* and *té* can take the classifiers *tôŋ* and *râŋ*, yielding *tôŋ ntû* and *râŋ té*. In the ordinary context, the tones of noun classifiers never trigger tone sandhi; however, in elaborate expressions and other co-compounds, they can, and \hat{x} is a tone-sandhi trigger. As a result, in this

context, /tôŋ ntû/ becomes [tôŋ ntɯ]. This change decisively affects the sequence of the conjuncts. Although $\hat{x} < \acute{x}$ along *R* and they are identical with regard to *T*, it is nevertheless true that $\acute{x} < \grave{x}$ along *T*. Thus, the predicted (and actual) output is tǎŋtété-*tóŋntɯ*:

		CORR-#σ ↔ #σ	WAX[T]	CORR-σ ↔ σ	WAX[R]
(227)	(a) tǎŋ _i ntɯ _i -tǎŋ _i tété		*	*!	
	(b) tǎŋntɯ _i -tǎŋtété _i	*!	*		
	(c) tǎŋ _i ntɯ _j -tǎŋ _i tété _j		*!*		
	(d) tǎŋ _i tété-tǎŋ _i ntɯ _i		*!	*	*
	(e) tǎŋtété _i -tǎŋntɯ _i	*!			*
	(f) tǎŋ _i tété _j -tǎŋ _i ntɯ _j		*		**

Note that, in this case, I do not address the exact mechanism by which tone sandhi is encoded in the grammar, a subject that is treated at some depth in Chapter 3 above. In this derivation, as with that for *fúatàì-vâŋntsí*, the best candidates are those where the initial syllables are co-indexed, satisfying CORR-#σ ↔ #σ at the expense of one violation of WAX[T]. Candidate (f), in which the order of *ntû-té* has been reversed, is the winner because it satisfies CORR-σ ↔ σ, in contrast to candidate (a), and does it without incurring an additional violation of WAX[T], in contrast to candidate (c).

It is significant that this set of candidates, where final-syllable differences along *T* beat out initial-syllable differences along *R*, run counter to the original generalization that we stated regarding coordinate compounds of this type, but are predicted to exist by the grammar that we developed to account for different phenomena. By its very nature—by the fact that it features competing but prioritized pressures determining the internal organization of linguistic structures—this case presents an interesting argument for Optimality Theoretic grammars. On the other hand, the nature of the scales that must be invoked seems inimical to certain assumptions that have been associated with Optimality Theory, the set of tendencies that Hale and Reiss (2000) have attacked as “substance abuse.” It seems impossible to say that the structures that emerge from this

grammar are better-formed in terms of their substantive, phonetic structure than the alternative outputs would be, at least in any universal or typological sense. Yet the very patterns they display and the criteria that must be invoked in explaining them explicitly and directly highlight the optimizing nature of these processes. The optimization is not driven by phonotactic well-formedness (as we suggested might be the case with Old Chinese co-compounds) or some cross-linguistically robust pattern (as is clearly the case in Jingpho, Tangkhul, and probably in the case of Lahu as well). Instead, these structures are optimized according to a set of abstract, language-specific structures—scales that are purely logical in their make-up, with no synchronic phonetic coherence at all.

5.6.3 Explaining the origin of the scales

It naturally follows that we should ask how the grammar of a language could come to include such an involved yet arbitrary set of relationships. It is also natural to ask whether the relationships truly are arbitrary, or whether they are coherent in a way that is easily missed if we look only at the phonetic exponence of the categories. It has already noted that the register scale *R*, which is motivated both diachronically and synchronically, forms an essential component of the ordering generalization. Scale *T*, too, appears to have a basis in facts that exist outside of this morphological construction.

As was discussed above, *R* corresponds to what was originally a contrast in onset voicing, with the low end of *R* representing what were originally ordinary voiced segments (voiced stops, fricatives, affricates, nasals, liquids, and glides) and the high end representing other onset consonants (voiceless [including voiceless aspirated] stops, fricatives, and affricates; “voiceless” nasals and liquids; preglottalized nasals, liquids, and glides). On comparative grounds, it appears that the effect of “*R*” on the ordering of co-compounds preceded the replacement of this voicing contrast with a tonal contrast since, as we will see, the same relationship is present in the Eastern Hmongic

(滇东方言 *diǎndōng fāngyán*) language Qe-Nao and the register split occurred independently in the several branches of Hmongic (leaving certain groups in e.g. Western Hmongic unaffected)²⁰. In other words, the original tendency was toward having a plain-voiced onset word initially, avoiding a plain-voiced onset word-internally, or preferring a marked-voiced (voiceless, voiceless aspirated, preglottalized) onset in the final syllable. As will be seen below, when the case of Qo-Siong is discussed, this particular preference cannot be universal, since it is exactly reversed in that language. However, there may be a general logic to it.

The scale *T* is somewhat more difficult to pin down. There is little evidence that would make clear the phonetic exponents of these categories at the time when the ordering effect came into existence, at least in terms of their pitch features. The one thing can be stated decisively about these categories at the time the ordering bias emerged regards their characteristic phonation type. It is notable that the “low register” subcategory of the historical B and C tones is reliably breathy across the whole Hmongic family (Niederer 1998; Andruski and Ratliff 2000). Likewise, the “low register” of the historical D tone is often creaky (Heimbach 1979; Ratliff 1992b), which is not surprising in light of the fact that D tone syllables come from syllables that ended, historically, in oral stop codas (Chang 1953; Purnell 1970; Wang 1994)²¹. While these differences are now found only in the low register, their existence is general in Hmongic, and must therefore predate the phonemic register split. Prior to this split, then, there must have been a voice-quality difference among the several tone categories. We might then pro-

²⁰The fact that such a split occurred independently in different branches of Hmong-Mien might seem surprising, were it not for the fact that the same type of change occurred in various branches of Tai-Kadai, Mon-Khmer, Sinitic, and Tibeto-Burman (Gedney 1947; Haudricourt 1954b; Matisoff 1973b, 1974, 1991; Li 1977).

²¹It ought to be noted, however, that not all Proto-Hmong-Mien syllables with final stops are reflected in Hmongic languages as syllables with a D tone. Syllables with PHmM *-k codas bear the C tone in PHm, and only those with *-p and *-t codas are reflected with the D tone in Hmongic.

pose of Hmong tones that A was historically modal-voiced, B and C were historically breathy, and D was historically creaky. If we apply this to our scale *T*, we get the following sequence:

(228) modal (A) < creaky (D) < breathy₁ (B) < breathy₂ (C)

It is immediately apparent that the syllables which were most preferred in the first position are those with modal phonation (not only in terms of the syllable as a whole, incidentally, but also in terms of the onset). In second position, breathy syllables are best, but barring those, creaky syllables are better than modal voiced syllables. If the pattern originally emerged in disyllabic compounds, and if the second member of the compounds was stressed, then this could be seen as the attraction of laryngeally marked structures to prosodically prominent positions. The association between stress and breathy or creaky voice is not unknown and the similar association between stress and voiceless (and especially, aspirated) onsets is also widespread²². This general model provides a good first approximation of the conditions and preferences that may have led to the emergence of this effect, and the accompanying scales. It does not, however, answer several important questions regarding both details (the relative relationship of B and C, the extension of patterns from two-syllable compounds to more complicated cases) or large issues (the initial status of the “preferences” that have been invoked and the mechanism by which these came to be encoded in the grammar).

A brief model of the development of ordering effects of this type may be presented as follows. There are certain stylistic preferences that guide the production of human speech and which exist above the level of “grammar” properly-so-called. Some of these seem to be universal (a preference for small things before big things, high vowels before low vowels, etc.) and some of these seem to be conditioned by facts about an individual

²²However, note that, in Lahu, aspirated consonants are preferred in initial position in coordinate compounds even though it is the final position which seems most likely to be prosodically prominent. This sort of principle does not seem to be universal in its application.

language (what laryngeal features are favored in particular positions). Since the internal organization of co-compounds is initially not fixed by the grammar, such stylistic preferences are given free play to manipulate co-compounds so that they are agreeable to these proclivities. However, by their nature, these preferences are not categorical, and may compete and interact with one another (in both systematic and non-systematic ways). This imposes a set of biases upon the data-set to which a language-learner is exposed. While these effects are still small, the learner may be able to tacitly recognize that the ordering principles are stylistic rather than grammatical; however, this will not stop them from settling on conventionalized forms that are constructed according to this set of competing stylistic desiderata. At length, learners are presented with data so biased that they start to seek out generalizations in it; if these generalizations are relatively involved—as they were in Hmong—the learner will attempt to recruit existing phonological scales (or failing that, propose new scales) in order to account for the graded distinctions that have emerged from the interactions of the original preferences. This is incorporated into the grammar, and as a result all new co-compounds are structured exactly according to this generalization, with a few exceptions still lurking in the lexicon from a time when the ordering tendencies were not yet grammaticalized. In the process of time, the phonetic generalizations upon which the original stylistic preferences may disappear. What remains, independent of its substantive roots, is a set of structural relationships that may be expressed in terms of scales and constraints that refer to them.

5.7 Co-Compounds in Qe-Nao

Hmong is not the only language which displays a tonally driven ordering effect of the type we have discussed. A very similar phenomenon, which turns out to have the same historical etiology, can be found in the Eastern Hmongic language, Qe-Nao. As shown

in Figure 5.1, Qe-Nao is quite distant from Hmong (*Chanqiandian*), but that they do belong to a common subgroup within Hmongic that is here called “Southern Hmongic.” Qe-Nao has six contrasting tonal categories: super high, high, low, super low, rising,

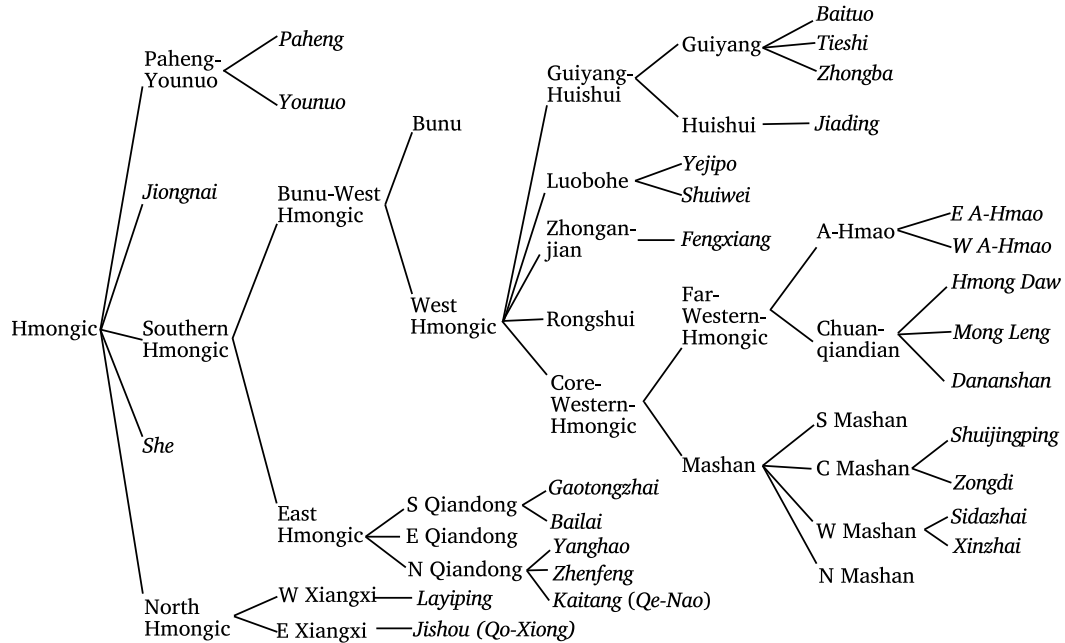


Figure 5.1: The Hmongic languages.

and falling. The distribution of these tones in 100 coordinate compounds (really, in “elaborate couplets”) is shown in (229):

(229) **Qe-Nao tone distribution**

T1 \ T2	ǎ	x	à	á	â	ǎ
ǎ (<i>super-high</i>)	1	1	6	11	5	11
x (<i>mid</i>)		1	11	4	1	14
à (<i>low</i>)			6	1	2	7
á (<i>high</i>)	1	2		2	2	6
â (<i>falling</i>)					—	3
ǎ (<i>rising</i>)			2			—

As can be seen by casual observation, there is an ordering generalization that holds for 95% of the cases encountered in this survey.

- | | |
|--|---|
| <p>(230) $\check{x} \prec x$</p> <p>a. ηo vě - tsâu ta
cloudy sky dark earth
'deep darkness'</p> | <p>(233) $\acute{x} \prec \hat{x}$</p> <p>a. $\check{h}au$ pí - $\check{h}au$ pô
purple EXPR purple EXPR
'purple (unattractive)'</p> <p>b. $qa\eta$ s^hé - $qa\eta$ ʎâ
boast words boast MORPH
'talk big'</p> |
| <p>(231) $x \prec \hat{x}$</p> <p>a. $tě$ lo - $tě$ pi
kick foot kick hand
'dance'</p> <p>b. faŋ - ɣàŋ
place village
'villages; places'</p> | <p>(234) $\hat{x} \prec \check{x}$</p> <p>a. tsêo t^hêo - tsêo k^hĩ
run about run MORPH
'to scamper about'</p> <p>b. $s^{h}ê\eta$ p^hêŋ - $s^{h}ê\eta$ k^hĩ
fell around feel MORPH
'to feel around indiscrimi-
nately'</p> |
| <p>(232) $\hat{x} \prec \acute{x}$</p> <p>a. $ha\eta$ kùŋ - $ha\eta$ kí
walk road walk road
'to walk'</p> | |

What is striking about the generalization we draw from these data is that, despite its formal similarity to the Hmong scale, the phonetic substance of the tones, or the tonal relationships within the sequence, seem to have little in common with their counterparts in Hmong. For Hmong, the pre-Theoretical scales was as follows:

- (235) \hat{x} (falling) \prec \acute{x} (high) \prec \check{x} (creaky) \prec \grave{x} (low) \prec \check{x}_1 (breathy₁) \prec \check{x} (rising) \prec \check{x}_2 (breathy₂) \prec x (mid)

For Qe-Nao, it is entirely different:

- (236) \check{x} (super-high) \prec x (mid) \prec \grave{x} (low) \prec \acute{x} (high) \prec \hat{x} (falling) \prec \check{x} (rising)

Despite this striking difference, the Qe-Nao (236) and the Mong Leng scale in (235) seem to share a common historical origin. If we align these two scales according to

historical categories, then the similarity between the two scales emerges:

(237) a. \hat{x} (A2) \prec x (A1) \prec \hat{x} (B2) \prec \acute{x} (B1) \prec \hat{x} (D2/D1) \prec \hat{x} (C2/C1)

(Qe-Nao)

b. \hat{x} (A2) \prec \acute{x} (A1) \prec \hat{x} (D1) \prec \grave{x} (D2) \prec \grave{x} (B2) \prec \hat{x} (B1) \prec \grave{x} (C2) \prec x (C1)

(Hmong)

Note that, as in the Mong Leng pattern, the “low” register tones (those marked with “2”) come before their “higher” register counterparts. Not only that, the sequence of the four historical tones (our scale *T*) is similar for both languages: $A \prec B \prec C$. The principle difference is the relative position of D. The dissimilarity between these two scales are exaggerated by two facts: First, tone B1 has merged with tone B2 (as \hat{x}) and tone C1 has merged with tone C2 (as x) in Qe-Nao. Second, the phonetic realizations of the tones have changed dramatically, erasing most of the surface correlation between the two tone systems.

5.8 Co-Compounds in Qo-Xiong

Moving out of the Southern Hmongic group (as delineated in Figure 5.1), it is possible to find still other examples of tonal ordering effects. These, however, seem to have had a different historical source, at least in part. The best-documented example of such an effect is the one described by (Yu 2004) for Jishou Miao (a dialect of Qo-Xiong). As in Qe-Nao, mergers have reduced the eight tonal categories we expect on historical grounds to six in Qo-Xiong. These tones are indicated here using Chao tone letters²³:

²³The choice to employ this notational standard here, rather than the systems of diacritics employed elsewhere, is a result of the awkwardness inherent in marking two contrasting falling tones without the use of special symbols. The Chao tone numbers have the added benefit of iconicity.

(238) **Qo-Xiong tones**

high fall.	low fall.	high	mid	rising	low
↘	↙	↑	┘	↗	┘

Yu (2004) demonstrates that these six categories can be arranged in a scale such that the sequence of (almost all) coordinate compounds whose conjuncts differ in tone can be predicted.

The distribution of tonal patterns in Qo-Xiong coordinate compounds is shown in (239). As is shown in the table, Yu (2004) found that 94.5% of the compounds in his sample followed his proposed ordering generalization:

(239) **Tone sequences in Qo-Xiong cocompounds**

T1\T2	↘	↙	↑	┘	↗	┘
↘ (<i>high falling</i>)	14	14	32	20	32	27
↙ (<i>low falling</i>)		7	22	8	14	10
↑ (<i>high</i>)		1	16	9	23	8
┘ (<i>mid</i>)	1	2	3	6	13	5
↗ (<i>rising</i>)	1		2	1	8	4
┘ (<i>low</i>)		2	3	1		7

As indicated, Yu arrives at the following scale:

(240) ↘ (*high falling*) < ↙ (*low falling*) < ↑ (*high*) < ┘ (*mid*) < ↗ (*rising*) < ┘ (*low*).

This precedence scale differs from the scales that were set up for Hmong and Qe-Nao in a number of respects, some superficial and some fundamental. Superficially, the arrangement of phonetic tone categories is not correlated with either of the earlier scales. On a deeper level, the arrangement of tones on this scale in terms of historical categories is rather different from that of the two earlier Hmongic languages. Stated in terms of historical tones, the arrangement of the Qo-Xiong scale is as follows:

(241) A1 < A2 < B1 < B2 < C1 < C2/D1/D2

While the scales in the other two languages are:

(242) a. A2 < A1 < B2 < B1 < D2/D1 < C2/C1 (Qe-Nao)

b. A2 < A1 < D1 < D2 < B2 < B1 < C2 < C1 (Hmong)

The most striking difference between the Qo-Xiong scales and those from Southern Hmongic is the fact that high register tones come before their low register congeners in Qo-Xiong, while they come after their low register equivalents in Southern Hmongic (e.g. Qe-Nao and Hmong).

This set of generalizations can be illustrated at all points through the following examples:

(243) HIGH FALLING TONE FIRST

a. **high falling** < **high falling**

i. pa¹ci¹ - pa¹tca¹
horn splinter
'pointed objects'

ii. qo¹pə¹ - qo¹təu¹
blanket ROOT
'bedroom articles'

b. **high falling** < **low falling**

i. a¹phu¹ -
grandfather
a¹ŋa¹
grandmother
'ancestors'

ii. tɕi¹ke¹ - tɕi¹wa¹
street village
'town'

c. **high falling** < **high**

i. qo¹tɕi¹ - qo¹ɕɛ¹
tripe liver
'heart; intention'

ii. ta¹ku¹ - ta¹ba¹
bug ant
'ant type insect'

d. **high falling** < **mid**

i. tɕi¹u¹ - tɕi¹pu¹
water slope
'water and land trans-
portation'

ii. ma¹ɕu¹ - ma¹ma¹
small fine
'small fine things'

e. **high falling** < **rising**

- i. ɲaŋ↓ hne↓ - ɲaŋ↓ nuŋ↑
time cold time hot
'winter and summer'
- ii. zɛi↓ - ʔe↑
vegetable rice
'food'

f. **high falling** < **low**

- i. qo↓pə↓ - qo↓tɕu↓
blanket seat
'bedroom articles'
- ii. hne↓ - nu↓
day day
'date'

(244) **LOW FALLING TONE FIRST**

a. **low falling** < **low falling**

- i. qo↓tɕu↓ - qo↓zɛ↓
generation age
'for generations'
- ii. qo↓tu↓ - qo↓ei↓
cudgel hammer
'a type of mallet'

b. **low falling** < **high**

- i. ku↑zɕu↓ - ku↑xaŋ↓
face up face down
'top and bottom'
- ii. pzu↑nu↓ - pzu↑dzu↓
house leaf house grass
'shack-type dwelling'

c. **low falling** < **mid**

- i. ɕuŋ↓ zɛ↓ -
Qo-Xiong remnant
ɕuŋ↓ tu↓
Qo-Xiong remnant
'bachelors'
- ii. zɛi↓ ɲe↓ -
vegetable water buffalo
zɛi↓ zu↓
vegetable ox
'cattle fodder'

d. **low falling** < **rising**

- i. pzu↑ ne↓ - pzu↑
house person house
qha↑
guest
'all the people of the
house'

- ii. qo↓ nuŋ↓ - qo↓ to↑
sickle hatchet
'tools of the knife and
hatchet type'

e. **low falling** < **low**

- i. u↓ ɲo↓ - u↓
water mouth water
mɕa↓
tongue
'saliva'
- ii. te↓wei↓ - te↓dzei↓
girl boy
'boys and girls'

(245) HIGH TONE FIRST

a. **high** < **high**

- i. pi¹lje¹ - pi¹po¹
STEM pimple
'pimples and the like'
- ii. qa¹ nu¹ - qa¹
stalk bean stalk soybean
tei¹

'bean stalks and the like'

b. **high** < **mid**

- i. qo¹pzu¹ - qo¹tə¹
house roaring fire
'house property'
- ii. qo¹ə¹ - qo¹pzu¹
clothing STEM
'dress'

c. **high** < **rising**

- i. ta¹təu¹ - ta¹ci¹
tiger STEM
'beasts of prey'
- ii. qo¹bei¹ - qo¹dzo¹
type of grain rice
'grain'

d. **high** < **low**

- i. qo¹tei¹ - qo¹tu¹
pit depression
'rugged land'
- ii. qo¹ce¹ - qo¹mu¹
breath STEM
'breath'

(246) MID TONE FIRST

a. **mid** < **mid**

- i. tci¹qə¹ - tci¹zə¹
stockade village
'settlement'
- ii. ne¹ ca¹ - ne¹
person stupid person
lja¹
stupid
'fool'

b. **mid** < **rising**

- i. qo¹ta¹ - qo¹te¹
plate bowl
'tableware'
- ii. ma¹pu¹ - ma¹dza¹
fear fear
'things that frighten people'

c. **mid** < **low**

- i. ne¹ -
father-in-law
mu¹
mother-in-law
'parents-in-law'
- ii. ne¹ təu¹ - ne¹
person done person
ku¹
sever
'windows and orphans'

(247) RISING TONE FIRST

‘power, force’

a. **rising** < **low**

- i. qo\ci[↑] - qo\nu[↓]
wind rain
‘wind and rain’

- ii. qo\kha[↑] - qo\zu[↓]
breath strength

(248) LOW TONE FIRST

a. **low** < **low**

- i. qo\te[↓] - qo\la[↓]
embankment field
‘embankments and fields’

This set of facts raises a question of considerable significance for the thesis presented here: Do the three instances of co-compound ordering effects in Hmongic languages (that have been examined here—there are many others) share a common origin, or are they independent developments? If they are independent, why are tonally-driven ordering constraints of the type seen here not more widespread? If they share a common origin, what is to account for the differences between them? Furthermore, what is the relationship between these ordering effects and those in Chinese languages?

As we have stated already, all of the Hmongic scales share two features in common, namely, the precedence relationships $A < B$ and $B < C$ (and thus, by transitivity, $A < C$). The relative position of the D tones is the first locus of variability, coming between A and B in Hmong, between B and C in Qe-Nao, and after C1 (confounded with C2) in Qo-Xiong. If we assume monogenesis, we must also assume that the position of D on the scale has changed somewhat in the various languages. If we suppose that the Hmong scale ($A < D < B < C$) was the same as the original scale, we can derive the Qo-Xiong scale simply by asserting that, when D1 and D2 merged with C2, syllables bearing these tones came to act like C2 in terms of ordering. However, this seems very improbable, since C2 is actually the least frequent of the tones (historically, at least) and it is difficult to imagine that it would “overwhelm” D1 and D2 in determining the relative sequence of co-compound conjuncts bearing the tone that resulted from this merger. This leaves either the $A < B < D < C$ pattern of Qe-Nao or the $A < B < C < D$

pattern of Qo-Xiong open as possibilities for the original ordering scheme, with the later scheme seeming somewhat more workable.

This ordering pattern actually matches the pattern seen in Chinese, since tones A, B, C, and D correspond (in loanwords from Middle Chinese) with the Chinese 平 *ping*, 上 *shang*, 去 *qu*, and 入 *ru* tones, and the Chinese scale is *ping* < *shang* < *qu* < *ru*. This introduces the possibility that this effect, and the scale that describes it, are shared by the two language families due to language contact. Under traditional assumptions, this fact might be attributed to the transmission of a feature from Chinese into Hmongic. After all, it may be observed that this phenomenon is attested in Chinese at a very early date (around 400 BC), which is believed to be prior to the date of tonogenesis in Chinese. On the other hand, there is increasing evidence that the contact relationship between Hmong-Mien languages and Chinese resulted in changes in both language groups (see Ballard (1986); Haudricourt and Strecker (1991), but see also Sagart (1995)). Indeed, it is possible that one of the factors that led to the development of tone in Chinese was contact with the highly tonal Hmong-Mien languages, which have no proven non-tonal relatives (indeed, no proven relatives) and thus for whom there is no evidence of time before tone. Under such a model, we might assume that this ordering effect originally developed in either Hmong-Mien or Chinese and was thus transmitted, in its A<B<C<D form, to the other language family. This constraint disappeared in much of the Hmong-Mien family, but was retained in two major branches of Hmongic, where it was transformed by some further changes.

There is at least one aspect of the ordering effect, however, that cannot be the result of shared inheritance or contact diffusion. That is the generalization captured by the scale *R* in Hmong—namely, the constraint that, other things being equal, mandates that tones in one register precede tones in the other. The reason that this cannot be a shared constraint is that it makes opposite demands in Northern Hmongic and South-

ern Hmongic. In Northern Hmongic, high-register tones must be ordered before their low-register equivalents; in Southern Hmongic, it is the low-register tones which must come first. In other words, even if the original pattern of constructing co-compounds according to tone was shared by Chinese, Qo-Xiong, Qe-Nao, and Hmong, the individual developments within these languages show that there was nothing magical about this genesis. Indeed, the Lahu case shows how tonal ordering effects can emerge, more or less on their own. It seems very likely that, given large enough corpora of co-compounds from tonal languages, many other low-level tonal effects on the internal organization of these compounds would be discovered. The only thing that is special about the Hmongic cases is that they have progressed beyond tendencies to become (nearly) deterministic grammatical processes.

5.9 The Significant of Coordinate Compounds for Phonology

As I have argued previously, it is significant that co-compound ordering effects, as grammatical processes, are not dependent upon the kinds of substantive biases that seem so essential to their emergence. A comparison of the Chinese, Hmongic, and Lahu cases make this clear. In Lahu, ordering is based on phonetically transparent criteria. Likewise, in the case of Chinese, it is possible that the motivations for ordering were based in natural phonetic preferences. However, these effects were not categorical; the tonal ordering effects that are strongest—that achieve the same level of robustness as the Jingpho and Tangkhul vowel-driven constraints—are those which are phonetically most opaque. An adequate theory of phonological grammar must be able to explain these facts.

Grammars are able to encode these generalizations because they possess representational devices capable of capturing them and formal machinery capable of operating on these representations. We have seen that phonological scales—when agnostic with

regard to phonetic substance—are able to capture these effects both elegantly and (as shown by the Hmong case) with considerable predictive power. It would be possible to model all of these effects with *ad hoc* mechanisms, but—in the end—all of these would have to be more stipulative than the mechanically simple proposal given here.

However, the very ability of these devices to model these ordering effects raises another question: Why do so many of the best examples of scale-driven effects in phonology involve the relative sequences of things (the sequence of segments according to sonority, the sequence of co-compound conjuncts according to vowel height or tone, etc.)? This fact is not as difficult to explain as it first appears. Most phonological facts for which phonologists attempt to account are the direct result of sound changes (“Jungphonologie”), a fact which has always been known, but which has gained recent prominence with the new wave of phonologists seeking historical explanations for synchronic patterns (Ohala 1974, 1981, 1993, 1995b,a; Hyman 2000; Dolbey and Hansson 1999a,b; Hansson 2001; Blevins and Garrett 1998, 2004; Blevins 2004; Hale 2000; Barnes 2002; Kavitskaya 2002; Mielke 2003, 2004a,b). Since, as this recent literature shows, the motivations for many sound changes are to be found outside of the grammar proper, it is reasonable to believe that the types of morphophonological patterns that are commonly observed represent a constrained set of the alternations that a human learner would be able to produce. Evidence for this proposition comes from a variety of sources, including the existence of so-called “crazy rules” (Bach and Harms 1972; Anderson 1981) and the learnability of arbitrary phonological alternations by both adults and infants (Pycha et al. 2003; Seidl and Buckley 2005) (but see also Wilson (2003)). This means that it is necessary to look to phenomena other than conventional morphophonological alternations in order to understand the true properties of the human capacity to learn and manipulate the sound patterns of language.

The ordering of coordinate compounds is a potentially phonological process that

involves sound-change only tangentially. Given this freedom from the inherently binary alternations that are typically created by sound changes (where α becomes β after X but remains α elsewhere) learners are able to employ their capacity to form n -ary generalizations to their full extent, discovering ambient patterns that accord with hierarchies rather than dichotomies. In other words, the presence of scalar generalizations in coordinate compounds is not a result of some intrinsic difference in the grammatical mechanism speakers use to grasp these patterns as opposed to morphophonological alternations; rather, the same formal tools that learners employ in constructing scales to order compounds could be used to construct scales for more conventional alternations. Chain shifts are one demonstration that this is the case. The reason that scalar effects are not more common in ordinary phonological alternations is quite simply that scales are not required to encode most of the patterns that history hands to the constructors of synchronic grammars.

Chapter 6

Logical Scales and Echo Reduplication

Woe unto them that call evil good,
and good evil;
That put darkness for light,
and light for darkness;
That put bitter for sweet,
and sweet for bitter!

Isaiah 5:20 (KJV)

In Chapter 5, we looked at two kinds of effects in coordinate compounds and echo reduplication—the ordering of conjuncts on phonological grounds and the avoidance of similarity between the two conjuncts. In that chapter, we looked at these phenomena largely from the perspective of co-compounding. In this chapter, by way of contrast, we will look at them from the standpoint of echo reduplication and show that these two types of constructions (co-compounds and echo reduplications) are not only related in many respects, but also that they present further confirmation of the claims that have been made thus far concerning phonological scales.

First, in §6.1, we will look at English echo reduplication, not because it is particularly enlightening with regard to the structure and nature of phonological scales, but because it provides a familiar context in which to introduce some of the other technology we will use in later analyses. Then we will apply the insight gained through the analysis of English echo-reduplication to echo reduplication in Jingpho in §6.1 and to Eastern A-Hmao in §6.3. The Eastern A-Hmao case is particularly interesting; I will show that it provides evidence for a vowel quality scale that is grounded in neither

height, backness, or roundness.

6.1 English

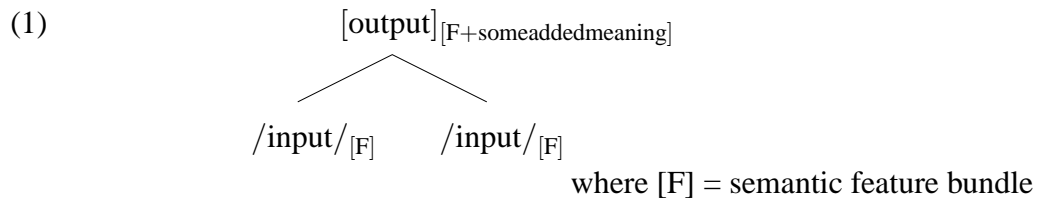
Just as there are ordering effects in coordinate compounds, there are ordering effects in echo-reduplication constructions. This idea was discussed briefly in §5.1.2.3 above, and it was mentioned that this fact can be seen even in English echo reduplication. Thus, in *hurry-scurry* and *handy-dandy*, the base form comes first and the modified form, second; however, in *hurly-burly* and *hubble-bubble*, it is the first conjunct that has been derived by some phonological process while the second conjunct is faithful to the underlying form. The same can be said for so-called ablaut-reduplication constructions. For example, the base forms in *click-clack*, *drizzle-drazzle*, *fiddle-faddle* and *jingle-jangle* are initial while those in *clitter-clatter*, *dilly-dally*, *crisscross*, and *wibble-wobble* are final. Where the base conjunct appears in this construction seems to be determined entirely by its phonological properties: it is positioned so as to fill the template.

6.1.1 Echo-reduplication and Morphological Doubling Theory

It is possible to look at this particular relationship in terms of Morphological Doubling Theory (MDT), a theory of reduplication in which the similarity between the base and reduplicant is captured not through copying or correspondence, but morphological identity (Inkelas and Zoll 2003, 2004, 2005). Differences between the two sisters (or, in the terminology I have used here, conjuncts) are the result of the different co-phonologies (co-grammars) associated with them. It is not difficult to see why this view of reduplication would be appealing in cases of this kind: we have already seen cases of coordinate compounding that display similar behaviors, namely conjuncts with high vowels being ordered before those with low vowels, regardless of the semantic content of the two conjuncts. For this and other reasons, echo reduplication looks a great deal

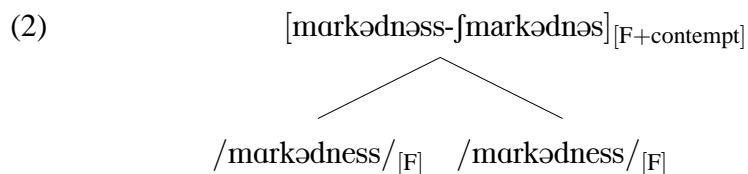
like coordinate compounding. MDT is interesting in that it treats all (or almost all) reduplication as something very much like compounding, a point that Inkelas and Zoll (2005:59–65) argue explicitly.

In MDT, a reduplication construction is conceived of as consisting of a mother node (the output of the construction) and two daughter nodes serving as the inputs to the construction. Take the following figure (adapted from Inkelas and Zoll (2005)):



In essence, a construction is a function—a function from all constructions of one type¹ to some subset of constructions of another type. However, it is not simply a function from phonological strings to phonological strings or semantic bundles to semantic bundles, a fact that is illustrated in the figure in (1), where both phonological and semantic relationships are represented. Rather, it takes as its input a collection that contains syntactic, semantic, and phonological information—a construction—and returns something of the same kind.

In this light, we could see *markedness-shmarkedness* as follows:



If we look at *shm*-reduplication in this way, then the construction takes two identical daughters, which are concatenated and subjected to the phonology that gives the second

¹The word *type* is used very loosely here. As we will see later, reduplication constructions may subcategorize for conjuncts having particular characteristics. Coordinate compounding constructions, likewise, may subcategorize for conjuncts that are semantically related in some way. In this case we would have to understand *type* as a set including all lexical items that meet particular criteria.

or string-internal constraints) to produce divergence between conjuncts. It is clear that string-internal correspondence plays a role in ordering effects, and these are important in echo reduplication, but it is not clear in all cases whether echo reduplication can be modeled using string-internal correspondence only or whether input-output relationships (e.g. input-output anti-identity) also play a role in these particular constructions.

Analytically, both types of processes will be discussed, for reasons that will become clear as we examine the several cases. Throughout the analyses, the MDT vision of reduplication as doubling will be assumed, to the exclusion of Base-Reduplicant Correspondence (McCarthy and Prince 1995; Alderete et al. 1996, 1999). Indeed, any important relationship that would be captured with base-reduplicant correspondence in BRCT (Base-Reduplicant Correspondence Theory) can be captured, for our purposes, by means of the more restrictive theory of string-internal correspondence. The notion of cophologies will also be important in these analyses. Indeed, they have been assumed throughout this dissertation, but here their specific invocation allows the phenomena to be described in a way that captures the morphological specificity of the phonological patterns involved.

6.1.2 A preliminary analysis of English echo-reduplication

We can exemplify one such analysis with the case of English echo-reduplication (as promised above). The generalization, in these constructions, is that the first conjunct tonic vowel is high (either /i/ or /ɪ/) and the second tonic vowel is low (/ɑ/ or /æ/). In the forms where there is an identifiable base, this base has one of these vowels as its tonic vowels, so at the very beginning we must stipulate that this construction imposes a subcategorization requirement such that its daughters may not have as their tonic vowels mid or rounded vowels. As we have observed on two previous occasions, the base (faithful conjunct, free form) may be either initial or final (although both the initial and final conjuncts may be free forms). This sequencing depends entirely upon whether

the free form contains a high or low vowel. The other conjunct is always a complement to the first, filling the position that it cannot. Both of the approaches to modeling echo reduplication that we have discussed above can be applied to this case.

Most straightforwardly, we could say that the reduplication construction takes two daughters, one of which is the unmodified free form and the other of which is a construction which has transformed that form into a suitable reduplicant by means of input-output anti-identity (our DIFF constraint or HIGHER constraint dominating an ENDMOST constraint) on a vowel-height scale. The jump from high vowel to low vowel is ensured by a constraint favoring entities at the extremes of the scale. The cophology of the reduplication construction then orders these two conjuncts, the two daughters, employing the same kind of mechanism that was explored in Chapter 5. Thus, given a word *drizzle* as an input, an “echo word” construction would evaluate as *drazzle*. The reduplication construction demands as inputs one “normal” word like *drizzle* and a semantically identical echo-word construction (in this case, *drazzle*). Its output, phonologically speaking, consists of the concatenation—in the phonologically appropriate order—of the phonological content of these two words.

For this particular case, an analysis of this sort has two disadvantages. The first is that two different constructions must impose similar subcategorization requirements: the reduplication construction demands daughters without mid or rounded vowels, and so does the “echo word” construction. The second redundancy regards reference to tonic vowels. Under this analysis, the “echo word” construction must make special reference to the tonic vowel to produce the appropriate output. The reduplication construction must then make reference to the tonic vowels of conjuncts to order them properly. It would be desirable if the subcategorization requirements could be stated only once and both the vowel change and ordering could be accomplished with a single positional reference.

This is, in fact, possible. On both theoretical and empirical grounds, we have motivated the constraint WAX. We employed it in the analysis of coordinate compounds, using it to explain the scalar ordering effects that are sometimes seen in these constructions. However, it is not difficult to see that a constraint of this type should also be able to motivate alternations. In cases where changing the sequence of correspondents is either impossible or cannot result in a satisfaction of WAX (that is, the correspondents are the same), the properties of individual segments (for example) may change in order to satisfy WAX, provided it outranks the faithfulness constraints that would militate against such changes. If we assume that the reduplication construction takes as its inputs two identical copies of the base, we have provided the setup for exactly this kind of situation. It may be, then, that both the sequence and the vowel disharmony are the result of a WAX constraint over a vowel height scale.

The analysis is very simple. Assume the following scale, H :

$$(4) \quad \{i, I, u, \upsilon\} < \{e, \varepsilon, o, \circ, \text{ə}\} < \{\text{æ}, \text{ɑ}\}$$

This implies the existence of a constraint WAX[H]. On independent grounds, we need a constraint CORR- $\acute{\sigma}$ – $\acute{\sigma}$ and we must assume that this constraint dominates WAX[H]. This ensures that WAX[H] cannot be satisfied vacuously through the absence of a correspondence relationship between the tonic vowels of the two conjuncts. Assuming that WAX[H] dominates PLATEAU[H], SAME[H], and any other faithfulness constraints that would prevent height alternations, this will yield outputs with a higher vowel in the first stressed syllable and a lower vowel in the second stressed syllable:

	CORR- $\acute{\sigma}$ ↔ $\acute{\sigma}$	WAX[H]	SAME[H]
(5) (a) dri _i zzle-dra _i zzle			*
(b) dri _i zzle-dre _i zzle			*
(c) dri _i zzle-dri _i zzle		*!	
(d) drizzle-drizzle	*!		

	CORR- \acute{o} \leftrightarrow \acute{o}	WAX[H]	SAME[H]	EXTREME[H]
(6) (a) dri _i zzle-dra _i zzle			*	
(b) dri _i zzle-dre _i zzle			*	*!
(c) dri _i zzle-dri _i zzle		*!		
(d) drizzle-drizzle	*!			

For Western American English, such a grammar adequately predicts the correct outputs in English ablaut reduplication constructions. However, it is worth noting that this analysis must be altered somewhat to account for other varieties of English, including the varieties that—judging from the orthography—existed at the time many of these forms were generated.

English ablaut reduplication, with its basis in vowel height, by no means the only example of a reduplication construction of this kind. A remarkably similar case is found in the Tibeto-Burman language Jingpho, the co-compounding construction of which was discussed above in §5.5.

6.2 Jingpho

As in the co-compounds, the first conjunct in Jingpho echo reduplication constructions always contains a tonic vowel at least as high as the corresponding vowel in the second conjunct. In fact, in reduplication this generalization is taken one step further to the extent that the tonic vowel of the first conjunct is always higher than the tonic vowel of the second conjunct.

6.2.1 Jingpho echo-reduplication and co-compounding

If the tonic vowel of the base is low (that is, /a/) the tonic vowel of the non-faithful conjunct will be /o/ and the non-faithful conjunct will be first (Dai 1990a; Dai and Xu 1992):

- (7) a. kumpháʔ ‘gift’ kumpho-kumpháʔ ‘every kind of gift’
 b. ləpʒá ‘fallen leaves’ ləpʒo-láphʒá ‘fallen leaves (all kinds)’

On the other hand, if the tonic vowel of the base is non-low, then the tonic vowel of the non-faithful conjunct will be /a/, and that conjunct will be ordered second (Dai 1990a; Dai and Xu 1992):

- (8) a. ákjîŋ ‘nervous’ ákjîŋ-ákâŋ ‘very nervous’
 b. kəmjíŋ ‘wrinkle’ kəmjíŋ-kəmján ‘in a wrinkled state’
 c. sə̀ùp ‘stuffy’ sə̀ùp-sə̀àp ‘sultry’
 d. ñhkjèŋ ‘crooked’ ñhkjèŋ-ñhkàŋ ‘very crooked’
 e. àkhjép ‘flat, thin piece’ àkhjép-àkháp ‘extremely fine’
 f. mə̀kjèp ‘glue’ mə̀kjèp-mə̀kàp ‘sticky’
 g. álóʔ ‘quarrel’ álóʔ-á aláʔ ‘quarrel’
 h. kàlóʔ ‘quarrel’ kàlóʔ-kà aláʔ ‘make a row’

This construction bears an obvious affinity to the coordinate compounding construction, given that both constructions involve the conjunction of two stems and that these stems are ordered according to the quality of the tonic vowels, with high-vowel stems being ordered before low-vowel stems. There is one significant difference, however, which would compel us to view the echo-reduplication construction as distinct from the co-compounding construction: In co-compounds, if the tonic vowels of the two conjuncts are identical in height, the order is indeterminate and both conjuncts are faithful to their underlying form. In the echo reduplication construction, however, this situation is resolved by changing the quality of the tonic vowel of one conjunct, allowing for a “contour” or “cline” across the two vowels relative to the vowel-height scale. This is crucial evidence that the paradigmatic relational constraint that drives ordering in these cases is WAX rather than NOWAX.

6.2.2 Explaining Jingpho echo-reduplication

What is the proper analysis of these reduplication constructions, however, and how does the cophology associated with co-compounds relate to that for echo-reduplications? When discussing co-compounds in §5.4, we concluded that the proper ranking for generating these forms is one in which $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ dominates $\text{WAX}[S]$. However, this ranking statement is clearly insufficiently specific. Take the example of *məsìn-səlum* ‘heart-heart’. Under this ranking, it is impossible to satisfy $\text{WAX}[S]$ by reordering the conjuncts, since both tonic vowels are equivalent in S . However, it would be possible to alter the quality of one of the vowels in order to satisfy this constraint:

	<i>məsìn-səlum</i>	$\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$	$\text{WAX}[S]$
(9)	(a) <i>məsì₇n-səl₇m</i>		*!
	(b) <i>səl₇m-məsè₇n</i>		
	(c) <i>məsì₇n-səl₀m</i>		
	(d) <i>məsìn-səlum</i>	*!	

This implies that, for the coordinate compounding construction, the constraint $\text{SAME}[S]$ is ranked above $\text{WAX}[S]$:

	<i>məsìn-səlum</i>	$\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$	$\text{SAME}[S]$	$\text{WAX}[S]$
(10)	(a) <i>məsì₇n-səl₇m</i>			
	(b) <i>səl₇m-məsì₇n</i>			
	(c) <i>səl₇m-məsè₇n</i>		*!	
	(d) <i>məsì₇n-səl₀m</i>		*!	
	(e) <i>məsìn-səlum</i>	*!		

The choice between *məsì₇n-səl₇m* and *səl₇m-məsì₇n* must ultimately be made based upon other factors.

This does not appear to be the right cophology for Jingpho echo reduplication because vowel qualities change in echo reduplication in order to satisfy $\text{WAX}[S]$. This

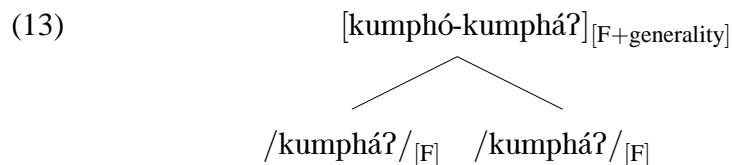
could never happen if SAME[S] were ranked above WAX[S]:

		CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	SAME[S]	WAX[S]
(11)	☞ (a) kumphá _i ?-kumphá _i ?			*
	(b) kumphó _i -kumphá _i ?		*!	
	(c) kumphá _i ?-kumphá _i ?	*!		

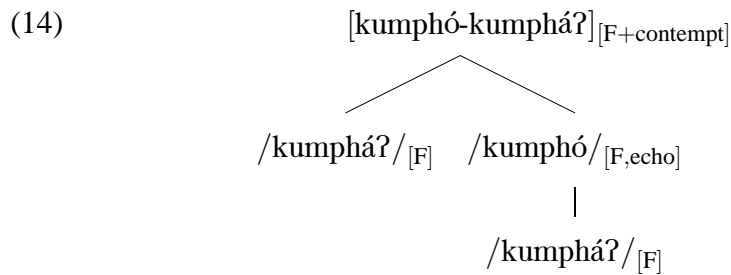
The correct output is admitted, however, by a grammar in which WAX[S] dominates SAME[S]:

		CORR- $\acute{\sigma} \leftrightarrow \acute{\sigma}$	WAX[S]	SAME[S]
(12)	(a) kumphá _i ?-kumphá _i ?		*!	
	☞ (b) kumphó _i -kumphá _i ?			*
	(c) kumphá _i ?-kumphá _i ?	*!		

These rankings, of course, assume that the reduplication construction has as its daughters two identical instances of the same stem and that any differences between the two conjuncts must be due to the cophonology associated with the reduplication construction. These relationships may be diagrammed as follows:



As mentioned earlier, there is a different way of looking at such constructions. We could actually assume that the cophonology associated with reduplication construction is exactly the same as that associated with the co-compounding construction and that the differences between the two conjuncts is due to differences in subcategorization rather than unfaithful mappings at the level of the reduplication construction. Diagrammatically, we can represent this construction as in (14):



There are advantages of both views of this construction, though ultimately it will be seen that the first is preferable. The second option has to recommend it the fact that only one cophonology is needed to accommodate both echo-reduplication and co-compounding. Since these two constructions are clearly related, this is a desirable result. This model also makes this type of reduplication a far-closer parallel to other types of reduplication discussed at length in Inkelas and Zoll (2005). The first option, however, captures certain generalizations that the second misses. First, the “shallow” model captures the fact that both ordering and alternation make reference to the relationship between the tonic vowels of the two conjuncts, a relationship that is easily captured through the constraint $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$. In order to compel the change in vowel quality in an input-output relationship, it would be necessary to invoke some special type of faithfulness (which may be needed in other places, but is unnecessary here). This type of analysis also allows us to avoid positing any input stems that do not exist as free forms. Of course, Inkelas and Zoll (2005) have argued persuasively that such bound intermediate forms are not uncommon in reduplication. However, the shallow analysis of echo-reduplication in Jingpho would allow us to avoid that issue altogether. The shallow analysis allows us to avoid multiplying entities—both in terms of lexical items and in terms of constructions.

An important question remains, however: what explains the nature of the apparent “fixed segmentism” associated with this construction? Why is the unfaithful vowel always /o/ or /a/? We may state a possible explanation of this fact informally. Suppose that, other things being equal, more sonorous vowels are preferable to less sonorous

vowels (in the head position, at least). Since the most sonorous vowel is /a/, then it is preferred as an output whenever this is feasible. This is feasible exactly when it produces a difference in the sonority of the tonic vowels of the two conjuncts. When the underlying vowel in the stem is /a/, then this cannot help and the output must be the next most sonorous option. In other words, this is a classical bounce-back effect. The next most sonorous vowels are the mid vowels /e/ and /o/. The most difficult question, in this case, is why /o/ is always chosen above /e/ when the input is not rounded and apparently not back. The stipulative response must be that /o/ is generally better in terms of markedness (relative to the cophology under discussion) than /e/.

We may formalize this as follows: there must be some constraint TOPMOST[S] that is dominated by SAME[S]. There are also low-ranked constraints *e and *o (standing in for a range of featural markedness constraints) with *e dominating *o. The ranking works as follows (counting violations of TOPMOST[S] only for tonic vowels):

(15)

	CORR- $\acute{\sigma}$ \leftrightarrow $\acute{\sigma}$	WAX[S]	SAME[S]	TOPMOST[S]	*e	*o
(a) kumphá _i ?-kumphá _i ?		*!				
■ ₃₈ (b) kumphó _i -kumphá _i ?			*	*		*
(c) kumphó _i -kumphá _i ?			*	*	*!	
(d) kumphú _i -kumphá _i ?			*!	**		
(e) kumphá _i ?-kumphá _i ?	*!					

(16)

	CORR- $\acute{\sigma}$ \leftrightarrow $\acute{\sigma}$	WAX[S]	SAME[S]	TOPMOST[S]	*e	*o
(a) kəmj _i n-kəmj _i n		*!		****		
(b) kəmj _i n-kəmjé _i n			*	***!	*	
(c) kəmj _i n-kəmjó _i n			*!	***		*
■ ₃₈ (d) kəmj _i n-kəmjá _i n			*	**		
(e) kəmj _i n-kəmj _i n	*!					

There are a number of details not resolved in this analysis, including the alternation between zero and glottal stop that is found in certain forms, or between palatalized and non-palatalized obstruents in others. However, this presentation should be sufficient to show both the fundamental unity of and the important differences between co-compounding and echo reduplication in Jingpho.

Were we to speak of constructions in terms of inheritance hierarchies, we could view both the co-compound (COCOMP) and echo reduplication ECHORED constructions to inherit from a common ancestor, which we might call DOUBLE. Inherent in double are a number of ranking declarations: $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma} \gg \text{WAX}[S]$, $\text{SAME}[S] \gg \text{TOPMOST}[T]$, $*e$, and $*o$, and possibly $*e \gg *o$. COCOMP adds a ranking statement to the effect: $\text{SAME}[S] \gg \text{WAX}[S]$. ECHORED, in contrast, adds the statement $\text{WAX}[S] \gg \text{SAME}[S]$. Thus, in adopting our “shallow analysis” of this construction, we do not really lose the generalization regarding the similarity between COCOMP and ECHORED. Instead, we locate this difference in terms of contradictory additions to a constraint hierarchy inherited from the same source.

It cannot escape notice that the Jingpho echo-reduplication construction is very much like its English analogue: in both constructions, the relative ordering of the two conjuncts is variable; and in both cases, the criterion for alternation and ordering appears to reside in the realm of vowel height or sonority. We will now look at a rather different case, namely that of Jingpho. It is similar to English (and, to a lesser extent, A-Hmao) in that the first conjunct always has a tonic vowel that is high. However, the quality of this vowel alternates and it is these alternations that provide additional evidence for the nature of phonological scales.

6.3 Eastern A-Hmao

6.3.1 Phonological patterning in Eastern A-Hmao reduplication

In Eastern A-Hmao, a dialect of A-Hmao that has been discussed in Chapters 3.2.2–4.2.2 above, there is an echo reduplication construction of some interest (already alluded to in §5.1.2.3). In this construction, described by Wang and Wang (1996) and Li (2003), there is a requirement that the tonic vowels of the two conjuncts differ in quality. The vowel of the second conjunct is the same as the underlying vowel. The vowel

of the first conjunct, however, must be a high vowel. There are four high vowels in Eastern A-Hmao, /i/, /y/, /w/ (actually [i]) and /u/, but only /i/ and /u/ appear in this position. The algorithm that determines which of these two vowels will appear is quite simple. If the rime of the tonic syllable contains no rounded segments underlyingly, the rime of the first conjunct will be /i/, as shown in (17):

(17) **No rounding in rime of tonic syllable**

- | | | | | |
|----|-----------------------------------|------------------|---|----------------------------|
| a. | pi ⁺ ntsi ⁺ | ‘butterfly’ | pi ⁻ ntsu ⁻ - pi ⁺ ntsi ⁺ | ‘butterflies, etc.’ |
| b. | ki ⁻ tə []] | ‘horn’ | ki ⁻ tu ⁻ - ki ⁻ tə []] | ‘horns of all kinds’ |
| c. | a ⁻ ma [\] | ‘eye’ | a ⁻ mu ⁻ - a [\] ma [\] | ‘eyes, mouth, and nose’ |
| d. | a ⁻ pha []] | ‘ingredients’ | a ⁻ phu ⁻ - a []] pha []] | ‘vegetables, tofu, etc.’ |
| e. | li ⁻ fai []] | ‘plowshare’ | li ⁻ fu ⁻ - li ⁻ fai []] | ‘plowshares, etc.’ |
| f. | ki ⁻ lau ⁻ | ‘strip of cloth’ | ki ⁻ lu ⁻ - ki ⁻ lau ⁻ | ‘strips of cloth and such’ |

If rime of the tonic vowel in the second conjunct is the single segment /u/, then its counterpart in the first conjunct will consist of the segment /i/, as show in (18):

(18) **Only /u/ in rime of tonic syllable**

- | | | | | |
|----|---------------------------------|------------------|---|-------------------------|
| a. | a ⁻ ndu [\] | ‘side’ | a ⁻ ndi ⁻ - a [\] ndu [\] | ‘thereabouts’ |
| b. | ki ⁺ tu ⁺ | ‘ridgepole tree’ | ki ⁻ ti ⁻ - ki ⁺ tu ⁺ | ‘ridgepole trees, etc.’ |
| c. | li ⁺ qu ⁻ | ‘bull’ | li ⁻ qi ⁻ - li ⁺ qu ⁻ | ‘bulls of all kinds’ |

However, if the segments underlying the tonic vowel include a rounded segment but are not /u/ alone, then the tonic vowel of the first conjunct can be either /i/ or /u/, with /i/ being somewhat more common, but with both forms being allowed in all cases (Wang and Wang 1996) as shown in (19):

(19) **Tonic syllable rime is not /-u/ and contains rounded segment**

a.	a ¹ tʂfiy ¹	‘cat’	a ¹ tʂfi ¹ - a ¹ tʂfiy ¹	‘cats’
			a ¹ tʂfiu ¹ - a ¹ tʂfiy ¹	‘id.’
b.	pi ¹ tʂo ¹	‘pimple’	pi ¹ tʂi ¹ - pi ¹ tʂo ¹	‘pimples, etc.’
			pi ¹ tʂu ¹ - pi ¹ tʂo ¹	‘id.’
c.	a ¹ phey ¹	‘basket’	a ¹ phi ¹ - a ¹ phey ¹	‘baskets of all kinds’
			a ¹ phu ¹ - a ¹ phey ¹	‘id.’
d.	a ¹ ntsau ¹	‘tree shadow’	a ¹ ntsu ¹ - a ¹ ntsau ¹	‘shade’
			a ¹ ntsi ¹ - a ¹ ntsau ¹	‘id.’
e.	a ¹ ndlfiau ¹	‘leaf’	a ¹ ndli ¹ - a ¹ ndlfiau ¹	‘leaves, etc.’
			a ¹ ndlu ¹ - a ¹ ndlfiau ¹	‘id.’
f.	a ¹ ndzfiu ¹	‘mouth’	a ¹ ndzu ¹ - a ¹ ndzfiu ¹	‘checks, noses, etc.’
			a ¹ ndzi ¹ - a ¹ ndzfiu ¹	‘id.’
g.	pi ¹ ndzau ¹	‘demon’	pi ¹ ndzu ¹ - pi ¹ ndzau ¹	‘spirits of all kinds’
			pi ¹ ndzi ¹ - pi ¹ ndzau ¹	‘id.’

The most immediately obvious fact about this pattern of reduplication is the avoidance of identity between the tonic vowels of the first and second conjuncts, where the vowel in the second conjunct is always faithful and that in the first conjunct is always unfaithful. The second obvious fact about reduplication in A-Hmao is that there is a three-way patterning of anti-identity relationships in the language. Rimes with no rounded element pattern one way, those with /-u/ pattern another way, and those with a rounded element but which are not /-u/ pattern in yet another way.

6.3.2 A-Hmao reduplication as a scalar phenomenon

It is easy to view this set of alternations as the result of a *PLATEAU constraint on some scale over the rimes of A-Hmao. The scale would have the following structure:

$$(20) \quad U = \{i, \text{ɪ}, \text{ə}, a, ai, au\} < \{y, o, ey, au\} < \{u\}$$

This scale is rather surprising in that, while it follows a phonetic dimension, it is not a dimension that is easily expressed in features and is certainly not a type of relationship that has been encoded in earlier scalar proposals. The basis of this scale seems to be relative similarity to /u/, with /u/ on one end, rimes completely lacking rounding on the other end, and intermediate cases in the middle. This case adds generality to the observation already made with tonal examples that phonological scales can be very flexible (and, in extreme cases, arbitrary).

The basic requirement imposed by the construction is that the two tonic vowels not be at the same point on the *S*. There are relatively low-ranked constraints that distinguish between /i/ and /u/ and it is the fact that these constraints are variably ranked (probably because there is relatively little evidence for their exact ranking) that produces the variation observed on the surface.

The fact that the tonic vowel in the first conjunct is always either /i/ or /u/ is due to the existence of another scale and a WAX constraint on that scale, this scale dividing “true” high vowels from non-high vowels. By “true” high vowels we refer to the vowels that are [+high] and in which [back] and [round] agree (meaning that they are not slightly centralized like /y/ and /ɪ/). The relevant scale, which we could call *S*, has the following structure:

$$(21) \quad S = \{a, ai, au, au, o, ey, \text{ə}, y, \text{ɪ}\} < \{i, u\}$$

This scale basically encodes the fact that /i/ and /u/ have a special status. Additionally, there must be a very high-ranked constraint IDENT-V_{σ]_ω} that prevents unfaithful mappings between the input and output features of word-final vowels. Let us now look at an implementation of this idea.

6.3.3 An analysis of A-Hmao echo reduplication

There are very interesting tonal patterns in Eastern A-Hmao reduplication, but since it is the segmental alternations that tell us the most about scales and constraints that refer to them, we will confine our analysis to this aspect of the construction. This choice has been made, in no small part, because an analysis of the tonal patterns requires making reference to the tone sandhi patterns of A-Hmao at a much more detailed level than is described in §4.2 above.

An analysis of the vocalic alternations in A-Hmao reduplication is both quite simple and strongly illustrative of the need for phonological scales and anti-identity constraints that make reference to them. The nucleus of this analysis has already been presented in 6.3.2. We start out with the following ranking statements:

- (22) a. $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma} \gg * \text{PLATEAU}[U]$,
b. $\text{IDENT-V}_{\sigma]_{\omega}} \gg * \text{PLATEAU}[U]$, $\text{WAX}[R]$, $*[+\text{round}]$, $*[-\text{round}]$
c. $* \text{PLATEAU}[U] \gg *[+\text{round}]$, $*[-\text{round}]$

We know, by the nature of the analysis, that $*[+\text{round}]$ and $*[-\text{round}]$ are variably ranked relative to one another. Because we have already observed how $\text{CORR-}\acute{\sigma} \leftrightarrow \acute{\sigma}$ functions in contexts of this type in earlier examples, the relative ranking of that constraint will be ignored in the tableaux. The relative ranking of $\text{WAX}[S]$ and $\text{PLATEAU}[U]$ cannot be determined based upon the available data. It is also the case that an additional constraint is needed to rule out certain special cases:

(23) **Tableau for pi-ntsi- ‘butterfly’**

pi-ntsi-	IDENT-V _{σ_iω}	*PLAT[U]	WAX[S]	*[+rnd]	*[-rnd]
⊃ (a) pi-ntsu _i - - pi-ntsi _i -			*	*	*
(b) pi-ntsi _i - - pi-ntsi _i -		*	*!		*
⊃ (c) pi-ntso _i - - pi-ntsi _i -			*	*	*
(d) pi-ntsu _i - - pi-ntso _i -	*!				
(e) pi-ntsu _i - - pi-ntsy _i -	*!				

In cases where it is not possible to change the “height” of the first vowel in order to satisfy WAX[S], it is no less harmonic to lower the vowel. This case needs to be ruled out by adding an additional constraint (actually already implied by the proposed inventory of scales, but not shown overtly in the ranking up to this point). This constraint, SAME[S], may be ranked at any point in the hierarchy (assuming it is ranked above such other unnamed constraints as DIFF[S]).

(24) **Tableau for pi-ntsi- ‘butterfly’**

pi-ntsi-	IDENT-V _{σ_iω}	*PLAT[U]	WAX[S]	SAME[S]	*[+rnd]	*[-rnd]
⊃ (a) pi-ntsu _i - - pi-ntsi _i -			*		*	*
(b) pi-ntsi _i - - pi-ntsi _i -		*	*!			*
(c) pi-ntso _i - - pi-ntsi _i -			*!	*	*	*
(d) pi-ntsu _i - - pi-ntso _i -	*!			*		
(e) pi-ntsu _i - - pi-ntsy _i -	*!			*		

Given this ranking, the unambiguous cases are easily generated, based upon logic that has already been explored:

(25) **Tableau for li-fai- ‘plowshare’**

li-fai-	IDENT-V _{σ_iω}	*PLAT[U]	WAX[S]	SAME[S]	*[+rnd]	*[-rnd]
⊃ (a) li-fu _i - - li-fai _i -				*	*	*
(b) li-fi _i - - li-fai _i -		*!		*		**
(c) li-fau _i - - li-fai _i -			*!		*	*
(d) li-fai _i - - li-fai _i -		*!	*			**

(26) **Tableau for a₁ndu₂ ‘side’**

a ₁ ndu ₂	IDENT-V _{σ₁ω}	*PLAT[U]	WAX[S]	SAME[S]	*[+rnd]	*[-rnd]
☞ (a) a ₁ ndi ₂ - a ₂ ndu ₁			*		*	*
(b) a ₁ ndu ₂ - a ₂ ndu ₁		*	*!		**	
(c) a ₁ ndy ₂ - a ₂ ndu ₁			*!	*	**	
(d) a ₁ ndi ₂ - a ₂ ndu ₁		*!	*		**	
(e) a ₁ ndu ₂ - a ₂ ndi ₁	*!		*		*	*

As we have said before, the variation between /i/ and /u/ is simply the result of variation in the ranking of the constraints *[+round] and *[-round] (for which there would be little evidence elsewhere in the language). The results of this can be exemplified as in (27) and (28). In (27), *[+round] dominates *[-round], meaning that /i/ is preferred over /u/ as the “fixed segment”:

(27) **Tableau for a₁tʃfi₂ ‘cat’**

a ₁ tʃfi ₂	IDENT-V _{σ₁ω}	*PLAT[U]	WAX[S]	SAME[S]	*[+rnd]	*[-rnd]
☞ (a) a ₁ tʃfi ₂ - a ₂ tʃfi ₁				*	*	*
(b) a ₁ tʃfu ₂ - a ₂ tʃfi ₁				*	*!	
(c) a ₁ tʃfi ₂ - a ₂ tʃfu ₁			*!		*	*
(d) a ₁ tʃfi ₂ - a ₂ tʃfi ₁		*!			**	

However, when *[-round] dominates *[+round], as in (28), /u/ is preferred above /i/:

(28) **Tableau for a₁tʃfi₂ ‘cat’**

a ₁ tʃfi ₂	IDENT-V _{σ₁ω}	*PLAT[U]	WAX[S]	SAME[S]	*[-rnd]	*[+rnd]
(a) a ₁ tʃfi ₂ - a ₂ tʃfi ₁				*	*!	*
☞ (b) a ₁ tʃfu ₂ - a ₂ tʃfi ₁				*		**
(c) a ₁ tʃfi ₂ - a ₂ tʃfu ₁			*!		*	*
(d) a ₁ tʃfi ₂ - a ₂ tʃfi ₁		*!				**

Thus, this grammar—with the assumption that variation is the result of coexisting grammars that differ in small ways—is able to model the whole set of patterns. It is able to do this using only mechanisms that have been motivated for other phenomena, adding to the body of evidence that human phonological grammars share important properties with Structural Optimality grammars.

6.4 Conclusion

Reduplication constructions are crucially important in the argument for logical scales and Structural Optimality in that they show that the same inventory of constraints used to model ordering effects (and even tone sandhi patterns like those in Jingpho) are also needed to account for segmental alternations in other contexts. This is important because it illustrates the logical scales are not in any way an ad hoc mechanism invoked only to account for a limited number of exceptional phenomena. Rather, the evidence for the presence of this type of relationship in the grammars of language is nearly ubiquitous and phonologists have failed to notice this, up to this point, simply because there was not yet a theoretical lens through which they could see them. This work generally, and no less this particular discussion of reduplication, are the first step to remedying this problem. It can now be seen that there are a whole range of phenomena that are driven by this grammatical mechanism.

In the case of Jingpho echo reduplication, and maybe English ablaut-reduplication as well, the same constraints produce alternations both in vowel quality and linear order, relating two functions of scales that have been observed earlier. These scales mirror the well-known sonority scale. However, the example from A-Hmao shows that scales can categorize types in far more flexible ways. In that particular case, there appears to be a scale over vowels or rimes dividing them according to their “roundness” or their similarity to /u/. Both of these cases illustrate the existence of anti-identity constraints, undermining versions of Optimality Theory that rest upon the idea that all unfaithful mappings are markedness reducing (see also §4.4). In this way, scalar effects in reduplication are a fitting conclusion to this exploration of logical scales in phonological representation and the novel way of looking at phonology that they encourage.

Chapter 7

Conclusion and Directions for Further Research

ἄλλη δόξα ἡλίου, καὶ ἄλλη δόξα
σελήνης, καὶ ἄλλη δόξα ἀστέρων:
ἀστὴρ γὰρ ἀστέρως διαφέρει ἐν δόξει.

I Corinthians 15:41
Paul of Tarsus

This chapter consists of a review of the findings of this study (in §7.1) followed by a presentation of several questions that have not been fully in this work but which are important avenues to be pursued in later investigations (in §7.2).

7.1 Conclusions

This study has had both a theoretical and an empirical component. While the theory presented in Chapter 2 might seem to be highly rarefied and far-removed from actual linguistic phenomena, the remaining chapters have demonstrated that this is not the case—the theoretical findings reported in that chapter have been shown to have important grounding in the empirical domain. I will now review the empirical phenomena that have been discussed in here, along with their theoretical relationship to Structural Optimality in §7.1, then discuss how these specific findings bear upon the whole endeavor of theoretical phonology.

7.1.1 Empirical phenomena and analyses

In general, it would be possible to divide the types of phonological phenomena discussed in this paper into two types: alternations and phonologically-driven morphotactics. In the first category could be placed the chain shifts and circle shifts discussed in Chapters 3 and 4 as well as some of the sound patterns in reduplication discussed in Chapter 6. Into the category of phonologically conditioned morphotactics would go the coordinate compound ordering effects collected and analyzed in Chapter 5 as well as some other effects in reduplication (in English and Jingpho) discussed in Chapter 6. This turns out to be a false dichotomy, however, because of the interesting overlap that exists between reduplication and coordinate compounding, and between the constraints that govern the sequence of conjuncts and the constraints that trigger segmental alternations. The case of Jingpho shows quite cogently how the line between reduplication and compounds—and between ordering and alternation—is a difficult and probably unnecessary line to draw. Instead, then, I will talk about the relevant phenomena one by one.

Of the phenomena treated by this dissertation, chain shifts are probably the best known and most widely discussed. They are the bread and butter of a certain class of historical linguists (Grimm's Law, the Great English Vowel Shift, and a huge set of similar shifts). However, they have presented some interesting problems for synchronic phonological analysis. Capturing chain shifts as unified phenomena, both within a particular shift, and across chain shifts as a class, has been difficult. This is true, in no small part, because chain shifts may come from a variety of different sources (see e.g. §3.2.1.1, §4.1.6) and because chain shifts that were originally grounded in phonetics may become quite arbitrary phonetically. Structural Optimality proposes that true chain shifts, whatever their origin, involve the traversal of a scale, driven by a single constraint (HIGHER). It thus delineates chain shifts from other types of counterfeeding opacity.

It also allows chain shifts to be treated as unitary processes while uniting chain shifts generally as a single *kind* of process. This is not all, though: in motivating the constraint HIGHER, chain shifts open the way to an interesting analysis of both circle shifts and co-compound ordering effects.

Circle shifts have received a great deal of discussion because of the fact that it has always been difficult to give a convincing analysis of them, whether in rule-based theories or constraint-based theories. A large number of investigators have tried to remove this type of phenomenon from the realm of phonology proper, whether by pointing out its (supposed) extremely confined distribution, arguing that it is the result of a morphological process of allomorph selection and not the result of the phonological grammar, or denying that such circles are phonological altogether. These three lines of argumentation have been made to appear more convincing by the fact that phonologists have all but ignored the diversity of circle shifts that do exist and concentrated almost exclusively on the tone circles in two closely related dialects of Southern Min (Chinese), one from Taiwan and one from Xiamen (on the mainland). Chapter 4 showed that circle shifts are actually far more diverse than this single-minded concentration on Xiamen would suggest, that circle shifts developed independently a number of times within Southern Min dialects and that they also exist in two geographically distant languages in Southeast Asia, namely A-Hmao and Jingpho. This is powerful evidence that such alternations cannot be dismissed as a marginal phenomenon with no bearing on phonological theory generally. In Structural Optimality, circle shifts are predicted to exist. In fact, a different permutation of the constraints needed to generate a “normal” chain shift will produce a circle shift of the type found in many Min dialects. This provides additional evidence for the necessity of the HIGHER constraint, which, in turn, gives a firmer basis for HIGHER’s string-internal analogue, namely, WAX. Somewhat different rankings characterize the bounce-back effects found in A-Hmao and Jingpho.

These patterns provide evidence for DIFF (a scale-referring, input-output, anti-identity constraint) and for NOWAX (a scale-referring string-internal directional faithfulness constraint).

The existence of the WAX constraint provides a ready explanation for why certain morphotactic generalizations in co-compounds should exist. The co-compounds that are studied at greatest depth in this work are of the “generalizing” type and are, as I demonstrate, part of a much broader continuum of phenomena that range from the *dvandva* compounds found in Indo-European, Dravidian, Japanese, and many other language families, to the echo reduplication constructions that are found widely in the languages of Eurasia (at least from the beginning of the historical era). Included in this continuum are also the irreversible binomial constructions that are so common in English and other European languages but which are found wide afield as well. All of these types of constructions are shown to allow some degree of phonological influence upon the linear sequences of their conjuncts. However, the strongest effects discovered in our survey (outside of the effects in echo-reduplication, which are treated in greater detail later in the study) are found in the “generalizing” co-compounds of East and Southeast Asia. These effects can be driven by vowel quality or tone (and, possibly, the laryngeal features of consonants). They can vary from the weak ordering effects in Chinese and Lahu, which are really just statistical biases, to the very strong ordering effects of Jingpho, Tangkhul, Hmong, Qe-Nao, and Qo-Xiong. A constraint equivalent to WAX appears to motivate these ordering generalizations. The case of Hmong (Mong Leng), treated in particular detail, demonstrates that there may be correspondence relationships between more than one pair of prominent positions in a single language and provides a striking confirmation for the kind of constraint-governed, string-internal correspondence proposed by earlier investigators like Walker (2000a,b), Hansson (2001), and Rose and Walker (2004). The upshot of this discussion is that

the phonological constraints governing phonologically conditioned morphotactic constraints are the same as those governing better-known phonological phenomena and that phonological theory cannot ignore the evidence provided for the structure of phonological representations and grammar by these effects.

This point is made even more forcefully when it is examined in light of various segmental alternations in reduplication constructions that echo the same set of principles seen in co-compound ordering effects. In the English and Jingpho cases, ordering effects are accompanied by the selection of segmental material that will satisfy WAX constraints. In the case of Eastern A-Hmao, the case for string-internal correspondence is made in a different context, as is the case for anti-identity. NOPLATEAU, the string-internal analogue of DIFF is motivated by an effect that demands difference, in terms of location on a three-point scale, between two corresponding vowels. Thus, these phenomena not only demonstrate the essential identity between string-internal constraints that determine the sequence of conjuncts in co-compounds and echo-reduplication constructions, they also show that the constraints that drive scalar input-output alternations drive output-oriented scalar alternations as well.

One of the great strengths of Structural Optimality, from an empirical point of view, is that it allows a wide variety of phonological phenomena that have typically been viewed as unrelated to be seen, instead, as the result of the operation of the same small set of phonological constraints in conjunction with a single, simple, representational device. However, Structural Optimality and the empirical phenomena that motivated it, have much deeper implications for phonological theory than this.

7.1.2 General theoretical implications

Structural Optimality is at once a return to some very traditional ideas about phonological theory and a radical rethinking of the concepts that have governed generative phonology. In a deeply ironic way, the most traditional aspect of the theory is also

the aspect that distinguishes it most sharply from the ideas that dominate phonology at the time of writing. As discussed in §1.1.1.1, the importance of phonetic substance and “naturalness” in phonological grammars has grown dramatically since the early 1970s, though acceptance of this line of thought has progressed gradually and acceptance of it has never been universal. This dissertation, instead, presents a markedly structuralist view of phonology in which the role of phonetic substance is far removed from the grammar itself and the relationship between phonological grammar and its phonetic instantiation is seen largely in terms of patterns in language learning and language change. The motivations for this position, however, are not primarily theoretical or philosophical: it is advocated in this study because it is empirically necessary. There are phonological phenomena that demand analyses in structural, rather than substantive, terms. This study is an attempt to provide a framework for analyses of this kind.

The representational primitives through which structural relationships of the logically-grounded type are established in this study are n -ary relationships, differing from the binary and primitive relations that have been most widely accepted in generative phonological theory. While earlier scholars have argued for scalar features, and while the evidence for such relationships continues to mount, few if any scholars have argued for the type of scalar relations—*independent of substance*—that have been advocated in this study. The somewhat counter-intuitive notion of scales as simple logical orderings rather than substance-grounded relations turns out to be a productive one.

The resulting machinery, while very simple, adds a great deal of power to the phonological grammar, and this will doubtless raise concerns in the minds of some phonologists, since it appears to undermine the findings and goals of much of contemporary phonological theory. After all, if phonological relations can exist independent of phonetic substance, many of the proposed universals in phonology can hardly be products of universal grammar as previously assumed. However, it is equally possi-

ble to view this kind of development as a liberation. Substantive tendencies can still be investigated, but can be investigated in terms of their causes in language acquisition and language change. The pursuit of grammatical universals has not been abandoned—instead, the phonologist is presented with a far more beautiful grammar where the generalizations and constraints are deeper and more profound. The effect of this move will be to tie phonological grammar far closer to morphology and syntax. This change, then, is not something to be feared but to be welcomed.

7.2 Remaining Questions and Directions for Further Research

A single study cannot address all of the questions that it raises, and this study has raised many. Some of these questions involve the empirical possibilities predicted by the theory and others, questions about the properties of the theory itself.

7.2.1 Bias toward Southeast Asian tone

One interesting fact about the case studies that have been presented in this work is that a disproportionate number of them have involved tone and that most of the remaining cases have centered around vowel quality. Also, of the tonal cases that motivate the use of scales, the great majority of those seen in this study are from a few language families in East and Southeast Asia. It is worthwhile to ask why these biases should exist. One simple explanation would be to look at the backgrounds and interests of the author. Since he is above all a Southeast Asianist and a tonologist, it is not unreasonable that he should draw examples from languages and phenomena with which he is familiar. By the same token, however, it might be supposed that his interest in these patterns is motivated precisely by the fact that they are present in the languages he has studied most closely. The question of the bias, then, should be investigated. If there is a bias, in the distribution of these phenomena, towards Southeast Asia, then it is important to

discover why.

Some answers to the bias question present themselves and should be examined more closely. It is not difficult to observe that it is far easier to conceive of vowels and tones in terms of phonetic continua than it is to conceive of consonant place and manner. Both vowels and tones, too, show a great deal of diachronic “mobility” in some language families. This mobility and sensitivity to perturbations, as suggested in this study, could contribute to the formation of chain shifts as well as the loss of grounding of such shifts and tonally-driven ordering effects. Since Southeast Asian tones, existing in large inventories in tightly-packed phonetic spaces, are especially prone to changes of this kind, it does not seem unreasonable that these effects should be most common in those languages. However, it is not easy to explain why certain languages, particularly Vietnamese those in the Daic family, may have relatively large tonal inventories with little evidence for tone sandhi and associated chain shifts, or even co-compound ordering effects.

In terms of the areal bias, a broader search for scalar phenomena may well uncover a much broader range of these patterns in different areas and language families. Indeed, there are a number of interesting scalar patterns found outside of Southeast Asia that have not been addressed in this study. The case of Esimbi vowel-height transfer (Stallcup 1980; Hyman 1988; Walker 1997) is especially interesting and should have great implications for the theory of Structural Optimality. Further research will doubtless turn up more cases of similar interest.

7.2.2 Logical scales and other features

It seems likely that the Esimbi case, which appears to show vowel features behaving as autosegments, will raise a question about Structural Optimality that has already been broached in the case of Jingpho tone sandhi: what is the relationship between logical scales and autosegmental features, or features in other geometrical relationships? On

the one hand, it appears that logical scales should replace conventional features altogether since they seem to perform a similar function. On the other hand, this would result in the loss of a large volume of work that has accumulated around autosegmental phonology and its descendants.

This issue may not be intractable. Features in earlier structuralist theories corresponded to categories. Starting in generative theories, and reaching full flower in autosegmental phonology, was the conception of features as actual entities—not characteristics of phonemes, but pieces from which segments were composed, and which could have an existence independent of individual segments. In Structural Optimality, it is possible to view representations as divided between entities and properties, where scales are (sets of) properties and autosegmental features, segments, syllables, and so on, are entities. Logically, scalar relationships are predicates which may or may not be true of entities, and entities are actual “things.” It may then be possible to derive binarity from privativity by means of scalar relationships, allowing the reduction of phonological features to one to one and reducing the apparent redundancy that exists between logical scales and the featural components of representations. This set of ideas deserves further thought and consideration.

7.2.3 Formal properties of Structural Optimality

Speaking more generally, it is important that the formal properties and implications of Structural Optimality be investigated at greater depth. Until recently, the formal properties of conventional Optimality Theory were not widely explored, but this line of research has accelerated, and now it is clearer what Optimality Theory really “means.” The same treatment ought to be given to Structural Optimality in order to determine what kinds of languages Structural Optimality can generate, what kinds it cannot generate, and how a Structural Optimality grammar and the accompanying generalizations can be learned algorithmically.

7.2.4 Scales outside phonology

Even knowing what Structural Optimality is like as a formal system does not answer a much larger question, which regards the relationship between scalar representations and constraints in phonology and other aspects of language—morphosyntax, semantics, and pragmatics. While scales and hierarchies of various kinds have been invoked in explaining phenomena in these domains, it is not immediately clear how the scales in Structural Optimality relate the n -ary relationships in other parts of the grammar. Once this is understood, it may be possible to establish the relation of phonology to other aspects of grammar more closely than before, and to see—in phonological phenomena—the play of general patterns in grammar in a way which was not previously possible. Should this research agenda bear fruit, it would be possible to place phonology far more securely in the realm of grammar than before.

Bibliography

- AHUM, VICTOR 1997a. *Tangkhul-Naga Grammar (A Study of Word Formation)*.
Mysore: Central Institute of Indian Languages.
URL <http://www.ciil-ebooks.net/html/tangkhul/CONTENTS.htm>
- 1997b. *Word formation in Tangkhul Naga*. Ph.D. thesis, Jawaharlal Nehru University.
- ALDERETE, JOHN D. 2001. “Dominance effects as transderivational anti-faithfulness”.
Phonology 18:201–253.
- ALDERETE, JOHN D., JILL BECKMAN, LAURA BENUA, AMALIA GNANADESIKAN,
JOHN MCCARTHY, AND SUZANNE URBANCYZK 1996. “Reduplication and segmental unmarkedness”. Ms, University of Massachusetts, Amherst [ROA 134-0996].
- 1999. “Reduplication with fixed segmentism”. *Linguistic Inquiry* 30(3):327–365.
- ANDERSON, STEPHEN R. 1981. “Why phonology isn’t ‘natural’”. *Linguistic Inquiry* 12(4):493–539.
- 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- ANDRUSKI, JEAN E. AND MARTHA RATLIFF 2000. “Use of phonation types in distinguishing tone: The case of Green Mong”. *Journal of the International Phonetic Association* 30:29–62.

- ARCHANGELI, DIANA B. 1999. "Introducing Optimality Theory". *Annual Review of Anthropology* 52:28–531.
- ARCHANGELI, DIANA B. AND DOUGLAS PULLEYBLANK 1994. *Grounded Phonology*. Cambridge, MA: MIT Press.
- BACH, EMMON AND ROBERT HARMS 1972. "How do languages get crazy rules?" In R. Stockwell and R. Macaulay (eds.), "Linguistic Change and Generative Theory", Bloomington, IN: Indiana University Press, 1–21.
- BAKOVIĆ, ERIC 2000. *Harmony, dominance, and control*. Ph.D. thesis, Rutgers University.
- BALLARD, WILLIAM LEWIS 1986. "The linguistic history of South China: Miao-Yao and southern dialects". In Graham Thurgood, James A. Matisoff, and David Bradley (eds.), "Linguistics of the Sino-Tibetan Area: The State of the Art", volume C-87 of *Pacific Linguistics*. Australian National University, 58–84.
- 1988. *The history and development of tonal systems and tonal alternations in south China*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- BAO ZHIMING 1999. *The Structure of Tone*. New York: Oxford University Press.
- BARNES, JONATHAN 2002. *Positional neutralization: a phonologization approach to typological patterns*. Ph.D. thesis, University of California, Berkeley.
- BARRIE, MICHAEL 2006. "Tone circles and contrast preservation". *Linguistic Inquiry* 37(1):131–141.
- BAXTER, WILLIAM 1992. *A Handbook of Old Chinese Phonology*. Berlin and New York: Mouton de Gruyter.

- BENOR, SARAH BUNIN AND ROGER LEVY 2006. "The chicken or the egg? a probabilistic analysis of English binomials". *Language* 82(2):233–278.
- BERKO, JEAN 1958. "The child's learning of English morphology". *Word* 14:150–177.
- BLEVINS, JULIETTE 2004. *Evolutionary Phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- BLEVINS, JULIETTE AND ANDREW GARRETT 1998. "The origins of consonant-vowel metathesis". *Language* 74:508–556.
- 2004. "The evolution of metathesis". In Bruce P. Hayes (ed.), "Phonetically-based Phonology", Cambridge: Cambridge University Press, 117–156.
- BLUST, ROBERT A. 1996a. "Low vowel dissimilation in Ere". *Oceanic Linguistics* 35:96–112.
- 1996b. "Low vowel dissimilation in Oceanic languages: An addendum". *Oceanic Linguistics* 35:305–309.
- BOERSMA, PAUL 1998. *Functional Phonology*. The Hague: Holland Academic Graphics.
- BOTTERO, FRANÇOISE 1996. "L'ordre des constituants dans les mots composés par coordination d'antonymes". *Cahiers de Linguistique Asie Orientale* 25(1):63–86.
- CHANG KUN 1953. "On the tone system of the Miao-Yao languages". *Language* 29:374–378.
- CHAO YUEN REN 1930. "A system of tone letters". *Le maître phonétique* 45:24–27.
- 1948. *Mandarin Primer*. Cambridge, MA: Harvard University Press.
- 1968. *Grammar of Spoken Chinese*. Berkeley: University of California Press.

- CHEN AIWEN 陈爱文 AND YU PING 于平 1979. “并列式双音词的字序 binglie shi shuangyinci de zixu [the tonal patterns of disyllabic coordinative compounds]”. *Zhongguo Yuwen* 1979(2):101–105.
- CHEN, MATTHEW Y. 1987. “The syntax of Xiamen tone sandhi”. *Phonology yearbook* 4:109–149.
- 2000. *Tone Sandhi: Patterns Across Chinese Dialects*. Cambridge: Cambridge University Press.
- CHENG, RICHARD L. 1968. “Tone sandhi in Taiwanese”. *Linguistics* 41:19–42.
- 1973. “Some notes on tone sandhi in Taiwanese”. *Linguistics* 100:5–25.
- CHENG, TSAI FA 1983. “The tonal features of Proto-South-Min”. *Papers in East Asian Languages* 1:66–90.
- CHOMSKY, NOAM AND MORRIS HALLE 1968. *The Sound Pattern of English*. New York: Harper & Row.
- CLEMENTS, G. N. 1990. “The role of the sonority cycle in core syllabification”. In John Kingston and Mary Beckman (eds.), “Papers in Laboratory Phonology”, Cambridge: Cambridge University Press, 283–333.
- CLEMENTS, GEORGE N. 1991. “Vowel height assimilation in Bantu languages”. *BLS* 17:25–64.
- CONTRERAS, HELES 1969. “Simplicity, descriptive adequacy, and binary features”. *Language* 45(1):1–8.
- COOPER, WILLIAM E. AND JOHN R. ROSS 1975. “World order”. In Robin E. Grossman, James L. San, and Timothy J. Vance (eds.), “Papers from the Parasession on Functionalism”, Chicago: Chicago Linguistic Society, 63–111.

- COURT, CHRISTOPHER 1985. “Observations on some cases of tone sandhi”. In Graham Thurgood, James A. Matisoff, and David Bradley (eds.), “Linguistics of the Sino-Tibetan Area: The State of the Art”, Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University, 125–137.
- BAUDOIN DE COURTENAY, JAN (1895 [1972]). “An attempt at a theory of phonetic alternations: a chapter from psychophonetics”. In Edward Stankiewicz (ed.), “A Baudoin de Cortenay anthology”, Bloomington: Indiana University Press, 144–212.
- DAI QINGXIA 戴庆厦 1990a. “景颇语并列结构复合词的音和谐 Jingpo yu binglie jiegou xiaheci de yuanyin he xie [on the vowels and agreement of Jingpho coordinate compounds]”. In Dai (1990d), 213–227.
- 1990b. “景颇语的声调 Jingpo yu de shengdiao [the tones of the Jingpho language]”. In Dai Qingxia 戴庆厦 (ed.), “藏面语族语言研究 Zang-Mian yuzu yuyan yanjiu [Research on Tibeto-Burman Languages]”, Yunnan Minzu Chubanshe, 199–212.
- 1990c. “藏面语族语言研究 Zang-Mian yu zu mou xie yuyan de jinjie dapei lu [collocation rules of syllables in some Tibeto-Burman languages]”. In Dai (1990d), 55–69.
- DAI QINGXIA 戴庆厦 (ed.) 1990d. 藏面语族语言研究 *Zang-Mian yuzu yuyan yanjiu [Research on Tibeto-Burman Languages]*. Kunming: Yunnan Minzu Chuban She.
- DAI QINGXIA 戴庆厦 AND XU XIJIAN 徐悉艰 1992. 景颇语语法 *Jingpo yu yufa [The Grammar of Kachin]*. Beijing: Zhongyang Minzu Xueyuan Chubanshe.
- DE LACY, PAUL 2002a. *The Formal Expression of Markedness*. Ph.D. thesis, University of Massachusetts, Amherst.

- 2002b. “The interaction of tone and stress in Optimality Theory”. *Phonology* 19:1–32.
- DICANIO, CHRISTIAN 2005. “Expressive alliteration in Mon and Khmer”. Unpublished qualifying paper, University of California, Berkeley.
- DINNSSEN, DANIEL AND JESSICA A. BARLOW 1998. “On the characterization of a chain shift in normal and delayed phonological acquisition”. *Journal of Child Language* 25:61–94.
- DOLBEY, ANDREW E. AND GUNNAR ÓLAFUR HANSSON 1999a. “Phonetic naturalness is not encoded in synchronic phonology”. Paper presented at HILP 4, Leiden, January 28–30, 1999.
- 1999b. “The source of naturalness in synchronic phonology”. *CLS* 35(I):59–70.
- DOWNER, GORDON B. 1967. “Tone-change and tone-shift in White Meo”. *BSOAS* 30:589–599.
- FLEMMING, EDWARD 1995. *Auditory Representations in Phonology*. Ph.D. thesis, UCLA.
- 2004. “Contrast and perceptual distinctiveness”. In Bruce P. Hayes, Robert Kirchner, and Donca Steriade (eds.), “Phonetically-based Phonology”, Cambridge: Cambridge University Press, 232–276.
- FOLEY, JAMES 1970. “Phonological distinctive features”. *Folia Linguistica* IV 1/2:87–92.
- 1972. “Reply to Cohen”. *CLS* 8:458–462.
- 1977. *Foundations of Theoretical Phonology*. London: Cambridge University Press.

- GEDNEY, WILLIAM 1947. *Indic Loanwords in Spoken Thai*. Ph.D. thesis, Yale University.
- GNANADESIKAN, AMALIA ELISABETH 1997. *Phonology with Ternary Scales*. Ph.D. thesis, University of Massachusetts, Amherst.
- GUTHRIE, MALCOLM 1968. “Notes on jzebi (Gabon)”. *Journal of African Languages* 7(2):101–129.
- HAAS, MARY R. 1964. *Thai-English Student’s Dictionary*. Stanford: Stanford University Press.
- HALE, MARK 2000. “Marshallese phonology, the phonetics-phonology interface and historical linguistics”. *The Linguistic Review* 17:241–257.
- HALE, MARK AND CHARLES REISS 2000. “‘Substance abuse’ and ‘dysfunctionalism’: Current trends in phonology”. *Linguistic Inquiry* 31(1):157–169.
- HANSSON, GUNNAR ÓLAFUR 2001. *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. thesis, University of California, Berkeley.
- HARRIS, ZELLIG S. 1951. *Structural Linguistics*. Chicago and London: University of Chicago Press.
- HAUDRICOURT, ANDRÉ-GEORGES 1954a. “Comment reconstruire le chinois archaïque”. *Word* 10:351–361.
- 1954b. “De l’origine des tons en vietnamien”. *Journal Asiatique* 242:69–82.
- HAUDRICOURT, ANDRÉ-GEORGES AND DAVID STRECKER 1991. “Hmong-Mien (Miao-Yao) loans in Chinese”. *International Journal of Chinese Studies* 77:335–341.

- HAYES, BRUCE AND DONCA STERIADE 2004. "Introduction: the phonetic bases of phonological markedness". In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.), "Phonetically based phonology", Cambridge: Cambridge University Press, 1–33.
- HAYES, BRUCE P. 1990. "Procompiled phrasal phonology". In Sharon Inkelas and Draga Zec (eds.), "The phonology-syntax connection", Chicago: CSLI and University of Chicago Press, 85–108.
- 1996. "Phonetically-Driven Phonology: The Role of Optimality Theory and Inductive Grounding". [ROA-158].
- HEIMBACH, EARNEST E. 1979. *White Hmong-English Dictionary*. Number 79 in Data Paper. Cornell University, Department of Asian Studies.
- HJELMSLEV, LOUIS 1961. *Prolegomena to a Theory of Language*. Madison: University of Wisconsin Press. Transl. Francis J. Whitfield.
- HOOVER, JOAN 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.
- HSIEH FENG-FAN 2004. "Tonal chain-shifts as anti-neutralization-induced tone sandhi". In Sudha Arunachalam, Tatjana Scheffler, Sandhya Sundaresan, and Joshua Tauberer (eds.), "Proceedings of the 28th Penn Linguistics Colloquium", volume 11.1 of *Penn Working Papers in Linguistics*. Philadelphia: Penn Linguistics Club, 99–112.
- HSIEH HSIN-I 1970. "The psychological reality of tone sandhi rules in Taiwanese". *CLS* 6:489–503.
- 1976. "On the unreality of some phonological rules". *Lingua* 38:1–19.

- HYMAN, LARRY M. 1977. "Phonologization". In A. Jullian (ed.), "Linguistic studies presented to Joseph H. Greenberg.", Saratoga: Anma Libri, 407–418.
- 1988. "Underspecification and vowel height transfer in Esimbi". *Phonology* 5(2):255–273.
- 2000. "The limits of phonetic determinism in phonology". In Elizabeth Hume and Keith Johnson (eds.), "The role of speech perception in phonology", San Diego: Academic Press, 141–185.
- INKELAS, SHARON, CEMIL ORHAN ORGUN, AND CHERYL ZOLL 1997. "Implications of lexical exceptions for the nature of grammar". In Iggy Roca (ed.), "Constraints and derivations in phonology", Oxford: Clarendon Press, 393–418.
- INKELAS, SHARON AND CHERYL ZOLL 2003. "Morphological doubling theory". Ms UC Berkeley/MIT.
- 2004. "Morphological Doubling Theory I: Evidence for morphological doubling in reduplication". In Bernhard Hurch (ed.), "Studies on Reduplication", Number 28 in Empirical approaches to language typology. Berlin: Mouton de Gruyter, 65–88.
- 2005. *Reduplication : Doubling in Morphology*. Cambridge: Cambridge University Press.
- ITÔ, JUNKO AND R. ARMIN MESTER 1999. "The phonological lexicon". In N. Tsujimura (ed.), "The Handbook of Japanese Linguistics", Malden, MA: Blackwell, 62–100.
- 1986. "The phonology of voicing in Japanese, theoretical consequences for morphological accessibility". *Linguistic Inquiry* 17(1):49–73.

- 1995. “Japanese phonology”. In John A. Goldsmith (ed.), “The Handbook of Phonological Theory”, Oxford: Blackwell, 817–838.
- JAKOBSON, ROMAN, GUNNAR FANT, AND MORRIS HALLE 1951. *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge, Massachusetts: MIT Press.
- JAKOBSON, ROMAN AND LINDA R. WAUGH 1979. *The Sound Shape of Language*. Bloomington, IN: Indiana University Press.
- JESPERSEN, OTTO 1904. *Lehrbuch der Phonetik*. Leipzig: Teubner.
- JOHNSON, MICHAEL 1999. “Tone and phonation in Western A–Hmao”. *SOAS Working Papers in Linguistics* 9:227–251.
- 2002. “The reconstruction of labial stop-consonant clusters in Proto-Far-Western Hmongic”. *Transactions of the Philological Society* 100(1):25–58.
- KAVITSKAYA, DARYA 2002. *Compensatory Lengthening*. Ph.D. thesis, University of California, Berkeley.
- KENSTOWICZ, MICHAEL AND CHARLES KISSEBERTH 1977. *Topics in Phonological Theory*. New York: Academic Press.
- 1979. *Generative Phonology*. New York: Academic Press.
- KIRCHNER, ROBERT 1995. “Going the distance: Synchronic chain shifts in optimality theory”. Ms. University of California, Los Angeles [ROA-66-0000].
- 1996. “Synchronic chain shifts in optimality theory”. *Linguistic Inquiry* 27(2):341–350.

- KRÄMER, MARTIN 2000. “Yucatec Maya vowel alternations. arguments for syntagmatic identity and positional faithfulness”. In “Two papers on constraint domains”, SFB 282 Working Papers, no. 116. Düsseldorf: Institut für Sprache and Information, 23–47.
- LAI, WIENCE 2002. “Jingpo yu shuang yinjie hecheng ci bian diao de youxuan jieshi [an optimality theory analysis of tone sandhi in Jingpho disyllabic compounds]”. Ms., Chinese University of Hong Kong.
- LEVIN, JULIETTE 1985. *A Metrical Theory of Syllabicity*. Ph.D. thesis, Massachusetts Institute of Technology.
- LI FANG-KUEI 1977. *A Handbook of Comparative Tai*. Manoa: The University of Hawai'i Press.
- LI JIE 李洁 2004. “拉祜语的并列结构复合词 Lahu yu de binglie jiegou xiahe ci [Lahu coordinate compounds]”. *Minzu Yuwen* 2004(4):27–30.
- LI XINKUI 李新魁 1994. 广东的方言 *Guangdong de fangyan [the dialects of Guangdong]*. Canton: Guangdong Renmin Chubanshe.
- LI YUNBING 李云兵 2003. “苗语的形态及其语义语法范畴 Miaoyu de xingtai ji qi yuyi yufa fanchou [morphology and semantic grammar categories in Miao language]”. *Minzu Yuwen* 2003(3):19–28.
- LIEN, CHINFA 1989. “Antonymous quadrinominals in Chinese”. *Journal of Chinese Linguistics* 17(2):263–306.
- LINDAU, MONA 1978. “Vowel features”. *Language* 54:541–563.
- LOMBARDI, LINDA 1999. “Positional faithfulness and voicing assimilation in optimality theory”. *Natural Language and Linguistic Theory* 17:267–302.

- ŁUBOWICZ, ANNA 2003. *Contrast Preservation in Phonological Mappings*. Ph.D. thesis, University of Massachusetts, Amherst.
- LYMAN, THOMAS AMIS 1974. *Dictionary of Mong Njua. A Miao (Meo) Language of Southeast Asia*. Mouton: Mouton.
- LYNCH, JOHN 2004. “Low vowel dissimilation in Vanuatu languages”. *Oceanic Linguistics* 35(2):359–409.
- MALKIEL, YAKOV 1959. “Studies in irreversible binomials”. *Lingua* 8(2):113–160.
- MAO ZONGWU 毛宗武 AND LI YUNBING 李云兵 1997. 巴亨语研究 *Baheng yu yanjiu* [A Study of the Baheng Language]. Shanghai: Shanghai yuandong chubanshe.
- MAO ZONGWU 毛宗武, MENG CHAOJI 蒙朝吉, AND ZHENG ZONGZE 郑宗泽 (eds.) 1982. 瑶族语言简直 *Yaozu yuyan jianzhi* [sketch of the languages of the Yao nationality]. Beijing: Minzu Chubanshe.
- MARCHAND, HANS 1969. *Categories and Types of Present-Day English Word-Formation*. München: C.H. Beck, 2nd edition.
- MATISOFF, JAMES A. 1973a. *The Grammar of Lahu*. Number 75 in University of California Publications in Linguistics. Berkeley: University of California Press.
- 1973b. “Tonogenesis in Southeast Asia”. In Larry Hyman (ed.), “Southern California Occasional Papers in Linguistics, No. 1”, volume 1. Los Angeles: The Linguistics Program, University of Southern California, 71–95.
- 1974. “The tones of Jingphaw and Lolo-Burmese: Common origin vs. independent development”. *Acta Linguistica Hafniensia* 15(2):153–212.

- 1989. *The Dictionary of Lahu*. Berkeley and Los Angeles: University of California Press.
- 1991. “Jiburish revisited: tonal splits and heterogenesis in Burmo-Naxi-Lolo checked syllables”. *Acta Orientalia* 52:91–114.
- in press. *English–Lahu Lexicon*. Berkeley and Los Angeles: University of California Press.
- MCCARTHY, JOHN J. 1999. “Sympathy and phonological opacity”. *Phonology* 16:331–399.
- 2002. “Comparative markedness”. Ms. University of Massachusetts, Amherst [ROA-489].
- MCCARTHY, JOHN J. AND ALAN PRINCE 1995. “Faithfulness and reduplicative identity”. In Jill Beckman, Laura Dickey, and Suzanne Urbanczyk (eds.), “University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory”, Amherst: GLSA, 249–384.
- MEI TSU-LIN 梅祖麟 1970. “Tones and prosody in Middle Chinese and the origin of the rising tone”. *Harvard Journal of Asiatic Studies* 30:86–100.
- MIELKE, JEFF 2003. “The diachronic influence of perception: evidence from Turkish”. *BLS* 29.
- 2004a. *The emergence of distinctive features*. Ph.D. thesis, The Ohio State University.
- 2004b. “Moving beyond innate features: a unified account of natural and unnatural classes”. *NELS* 35.

- MOHANAN, KARUVANNUR PURHANVEETIL 1986. *The Theory of Lexical Phonology*. Dordrecht, Holland: D. Reidel.
- MORETON, ELLIOT 2004a. “A compendium of synchronic chain shifts”. Ms. University of North Carolina, Chapel Hill.
- 2004b. “Non-computable functions in optimality theory”. In John J. McCarthy (ed.), “Optimality Theory in Phonology”, Oxford: Blackwell, 141–164.
- MORETON, ELLIOT AND PAUL SMOLENSKY 2002. “Typological consequences of local constraint conjunction”. In Line Mikkelsen and Christopher Potts (eds.), “WCCFL 21”, Somerville, MA: Cascadilla Press.
- MORTENSEN, DAVID 2003. “Abstract scales in phonology”. Ms. University of California, Berkeley.
- 2005. “The development of tone sandhi in Western Hmongic: an abstract approach to diachronic morphophonology”. Ms. University of California, Berkeley.
- NEVINS, ANDREW AND BERT VAUX 2003. “Metalinguistic, schmetalinguistic: The phonology of schm- reduplication”. Paper presented at the Chicago Linguistics Society.
- NIEDERER, BARBARA 1998. *Les langues Hmong-Mjen (Miáo-Yáo)*. Number 07 in Lincom Studies in Asian Linguistics. München: Lincom Europa.
- NORMAN, JERRY 1988. *Chinese*. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- OHALA, JOHN J. 1974. “Experimental historical phonology”. In John M. Anderson and Charles Jones (eds.), “Historical linguistics”, volume 2. Amsterdam: North-Holland, 353–389.

- 1981. “The listener as a source of sound change”. In C. Masek, R. A. Hendrick, and M. F. Miller (eds.), “Papers from the Parasession on Language and Behavior”, Chicago: Chicago Linguistic Society.
 - 1990. “Alternatives to the sonority hierarchy for explaining the shape of morphemes”. In “Papers from the Parasession on the Syllable.”, Chicago: Chicago Linguistic Society, 319–338.
 - 1993. “The phonetics of sound change”. In Charles Jones (ed.), “Historical Linguistics: Problems and Perspectives”, London: Longman, 237–278.
 - 1994a. “Hierarchies of environments for sound variation; plus implications for ‘neutral’ vowels in vowel harmony”. *Acta Linguistica Hafniensia* 27:371–382.
 - 1994b. “Towards a universal, phonetically-based, theory of vowel harmony”. In “Proceedings of the International Conference on Spoken Language Processing (ICSLP)”, Yokohama, Japan.
 - 1995a. “Experimental phonology”. In John A. Goldsmith (ed.), “A Handbook of Phonological Theory”, Oxford: Blackwell, 712–722.
 - 1995b. “The phonetics of phonology”. In V. Hazan, D. Huber, and J. Llisterri (eds.), “European studies in phonetics and speech communication”, Utrecht: OTS Publications, 85–89.
- OLSEN, SUSAN 2000. “Copulative compounds : a closer look at the interface between syntax and morphology”. *Yearbook of morphology* :279–320.
- ORGUN, ORHAN AND RONALD SPROUSE 1999. “From MPARSE to CONTROL: deriving ungrammaticality”. *Phonology* 16:191–224.

- OURN, NOEURNG AND JOHN HAIMAN 2000. “Symmetrical compounds in Khmer”. *Studies in Language* 24(3):483–514.
- PAN, YUANEN AND CUIYIN CAO 1972. “Four-syllable coordinative constructions in the Miao languages of eastern Kweichow”. In Herbert C. Purnell (ed.), “Miao and Yao Linguistic Studies”, Number 88 in Data Papers. Ithaca, New York: Cornell University Southeast Asia Program, 211–234.
- PARKER, STEVE 2002. *Quantifying the Sonority Hierarchy*. Ph.D. thesis, University of Massachusetts, Amherst.
- POSTAL, PAUL M. 1968. *Phonological Theory*. New York: Harper & Row.
- PRINCE, ALAN AND PAUL SMOLENSKY 1993. “Optimality theory: Constraint interaction in generative grammar”. Technical Report 1, Rutgers University Center for Cognitive Science.
- PRZEZDZIECKI, MAREK A. 2005. *Vowel Harmony and Coarticulation in Three Dialects of Yoruba: Phonetics Determining Phonology*. Ph.D. thesis, Cornell University.
- PULLEYBLANK, EDWIN G. 1962. “The consonantal system of Old Chinese”. *Asia Major* 9:58–144, 206–265.
- 1973. “Some further evidence regarding Old Chinese *-s and its time of disappearance”. *Bulletin of the School of Oriental and African Studies* 36(2):368–373.
- PURNELL, HERBERT C. 1970. *Toward and Reconstruction of Proto-Miao-Yao*. Ph.D. thesis, Cornell University.
- PYCHA, ANNE, PAWEL NOWAK, EURIE SHIN, AND RYAN SHOSTED 2003. “Phonological rule-learning and its implications for a theory of vowel harmony”. In

- G. Garding and M. Tsujimura (eds.), “West Coast Conference on Formal Linguistics”, volume 22. Cambridge: Cascadilla Press.
- QUAN CHAOQIANG 全病善 1990. “论试汉，朝并列复合名词字序的共同点 Luan-shi Han, Chao binglie xiahe mingci zixu de gongtong dian [on common grounds of the word orders of coordinate compounds of Chinese and Korean]”. *Minzu Yuwen* 1990(6).
- RATLIFF, MARTHA 1992a. “The development of nominal/non-nominal class marking by tone in Shimen Hmong”. *BLS* 17:267–282.
- 1992b. *Meaningful Tone: A Study of Tonal Morphology in Compounds, Form Classes, and Expressive Phrases in White Hmong*. Number 27 in Monograph Series on Southeast Asia. DeKalb Illinois: Northern Illinois University Center for Southeast Asian Studies.
- RIVAS, A. M. 1977. “Hierarchical classes of features in binary-feature phonology”. *NELS* 8.
- ROSE, SHARON AND RACHEL WALKER 2004. “A typology of consonant agreement as correspondence”. *Language* 80(3):475–531.
- ROSS, JOHN R. 1967. *Constraints on variables in syntax*. Ph.D. thesis, Massachusetts Institute of Technology.
- SAGART, LAURENT 1995. “Chinese “buy” and “sell’ and the direction of borrowings between Chinese and Hmong-Mien”. *T’oung Pao* LXXXI(4–5):328–342.
- 1999. *The Roots of Old Chinese*. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam: John Benjamins.

- DE SAUSSURE, FERDINAND 1916. *Cours de linguistique générale*. Lausanne and Paris: Payot.
- SCHILLER, ERIC 1999. “Reduplication in Khmer morphology”. Ms. Linguistics Unlimited.
- SCHUH, RUSSELL G. 1978. “Tone rules”. In Victoria Fromkin (ed.), “Tone: a linguistic survey”, New York: Academic Press, 221–256.
- SEIDL, AMANDA AND EUGENE BUCKLEY 2005. “On the learning of arbitrary phonological rules”. *Language Learning and Development* 1(3&4):289–316.
- SELKIRK, ELISABETH O. 1984. “On the major class features and syllable theory”. In M. Aronoff and R. T. Oehrle (eds.), “Language Sound Structure: Studies in Phonology Dedicated to Morris Halle by his Teacher and Students”, Cambridge, Massachusetts: MIT Press, 107–113.
- SIEVERS, EDUARD 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Härtel.
- SIMPSON, ANDREW 2005. “Classifiers and DP structure in Southeast Asia”. In Guglielmo Cinque and Richard Kayne (eds.), “The Oxford Handbook of Comparative Syntax”, Oxford: Oxford University Press, 806–836.
- SIO, JOANNA UT-SEONG 2006. *Modification and reference in the Chinese nominal*. Utrecht: Netherlands Graduate School of Linguistics.
- STAHLKE, HERBERT F. W. 1977. “Some problems with binary features for tone”. *International Journal of American Linguistics* 43(1):1–10.
- STALLCUP, KENNETH L. 1980. “A brief account of nominal prefixes and vowel harmony in Esimbi”. In Luc Boquiaux (ed.), “L ’expansion bantoue”, volume 2. Paris: Société d’Etudes Linguistiques et Anthropologiques de France, 435–441.

- STALLINGS, LYNNE M., MARYELLEN C. MACDONALD, AND PADRAIG G. O'SEAGHDHA 1998. "Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift." *Journal of Memory and Language* 39:392–417.
- STAMPE, DAVID 1973. *A dissertation on natural phonology*. Ph.D. thesis, University of Chicago.
- STERIADE, DONCA 1982. *Greek Prosodies and the Nature of Syllabification*. Ph.D. thesis, Massachusetts Institute of Technology.
- 1997. "Phonetics in phonology: the case of laryngeal neutralization". Ms., UCLA.
- 2002. "The Phonology of Perceptibility Effects: the P-map and its consequences for constraint organization". Ms. UCLA.
- THOMPSON, LAURENCE C. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.
- TING PANG-HSIN 丁邦新 1969. "國語中雙音節並列語兩成分間的聲調關係 guoyu zhong shuang yinjie binglie yu liang chengfen jian de shengdiao guanxi (tonal relationship between the two constituents of the disyllabic coordinate constructions in Mandarin)". *Bulletin of the Institute of History and Philology, Academia Sinica* 39:155–173.
- 1975. "論語、孟子、及詩經中並列語成分之間的聲調關係 Lunyu, Mengzi, ji Shijing zhong binglie yu chengfen zhi jian de shengdiao guanxi [tonal relationship between the two constituent of the coordinate construction in the Analects, the Meng-tze, and the Book of Odes]". *Bulletin of the Institute of History and Philology, Academia Sinica* 47(1):17–52.

- TINNEY, STEVE (ed.) 2006. *The Pennsylvania Sumerian Dictionary*. Philadelphia: University of Pennsylvania Museum of Anthropology and Archaeology.
URL <http://enlil.museum.upenn.edu/epsd/>
- TSAY, JANE AND JAMES MYERS 1996. “Taiwanese tone sandhi as allomorph selection”. In Jan Johnson, Matthew L. Juge, and Jeri L. Moxley (eds.), “BLS”, volume 22. Berkeley, California: Berkeley Linguistics Society.
- VENNEMANN, THEO 1972. “On the theory of syllabic phonology”. *Linguistische Berichte* 18(1):1–18.
- WALKER, RACHEL 1997. “Faith and markedness in Esimbi feature transfer”. In R. Walker, Katayama M., and Karvonen D. (eds.), “Phonology at Santa Cruz 5”, Santa Cruz, California: University of California, Santa Cruz, Department of Linguistics, 103–115.
- 2000a. “Long distance consonantal identity effects”. *WCCFL* 19:532–545.
- 2000b. “Yaka nasal harmony: Spreading or segmental correspondence”. Paper presented at BLS 26.
- WANG DEGUANG 王德光 1986. “威宁苗语话语材料 Weining Miaoyu huayu cailiao [Language material of the Miao language in Weining]”. *Minzu Yuwen* 1986:69–80.
- WANG FUSHI 王辅世 1979. 苗语方言声韵母比较 *Miaoyu fangyan sheng yun mu bijiao* [Comparison of the tones, rhymes, and onsets of Miao dialects.]. Beijing: Zhongguo Shehui Kexue Yuan Minzu Yanjiushi.
- 1985. 苗语简志 *Miao yu jianzhi* [Sketch grammar of the Miao language]. Beijing: Minzu Chubanshe.

- 1994. 苗语古音构拟 *Miaoyu guyin gouni [Reconstruction of Proto-Miao]*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- WANG FUSHI 王辅世 AND WANG DEGUANG 王德光 1986. “贵州威宁苗语声调 Guizhou Weining Miaoyu shengdiao”. In Fu Maoji 傅懋勣 (ed.), “中国民族语言论文及 Zhongguo Minzu Yuyan Lunwen Ji [Collect of Papers on Chinese National Minority Languages]”, Chengdu: Sichuan Minzu Chubanshe, 91–134.
- 1996. “贵州威宁苗语带前加成分的双音节名词的形态变化 Guizhou Weining Miaoyu dai qianjia chengfen de shuangyin jie mingci de xingtai bianhua”. *Minzu Yuwen* 1996(1):34–36.
- WANG, WILLIAM S-Y 1967. “Phonological features of tone”. *International Journal of American Linguistics* 33(2):93–105.
- WATSON, RICHARD L. 1966. *Reduplication in Pacoh*. Master’s thesis, The Hartford Seminary Foundation.
- WILLIAMSON, KAY 1977. “Multivalued features for consonants”. *Language* 53(4):843–871.
- WILSON, COLIN 2003. “Experimental investigation of phonological naturalness”. *WCCFL* 35:101–114.
- XIAN SONGKUI 鲜松奎 1990. “贵州紫云水井坪苗语和望谟新寨苗语的连读变调 Guizhou Ziyun Shuijingping Miaoyu he Wangmo Xinzhai Miaoyu de liandu biandiao [tone sandhi in the Miao speech of Shuijingping, Ziyun and the Miao speech of Xinzhai, Wangmo in Guizhou]”. *Minzu Yuwen* 1990(3):71–77.
- XIONG, JAY 2005. *Lus Hmoob Txhais (English-Hmong Dictionary)*. Manitowoc, Wisconsin: HmongDictionary.com, 2nd edition.

- XIONG, LANG, JOUA XIONG, AND NAO LENG XIONG 1992. *English–Mong–English Dictionary (Phoo Txhais Lug Aakiv–Mong–Aakiv)*. Milwaukee: Xiong Partnership Publications (self published).
- XU YI 1997. “Contextual tonal variations in Mandarin”. *Journal of Phonetics* 25:61–83.
- YIP, MOIRA 1980. *The tonal phonology of Chinese*. Ph.D. thesis, MIT.
- 2002. *Tone*. Cambridge: Cambridge University Press.
- YU, ALAN 2002. “Understanding infixes as infixes”. Paper presented at the Second North American Phonology Conference, Montréal.
- YU JINZHI 余金枝 2004. “吉首矮寨苗语并列复合名词的结构和声调特征 Jishou Aizhai Miaoyu binglie xiahe mingci de jiegou he shengdiao tezheng [the structural and tonal particulars of the coordinate compound nouns of the Miao language of Jishou Aizhai]”. *Minzu Yuwen* 2004(1):27–29.
- ZHANG JIE 2006. “Wug-testing the “tone circle” in Taiwanese”. Paper presented at WCCFL 25.
- ZURAW, KIE 2002. “Aggressive reduplication”. *Phonology* 19:395–439.