Nothing to Lose but Their Chains:
Rethinking Vocalic Chain Shifting

A thesis submitted
by
Bridget Diane Samuels
to
the Department of Linguistics
in partial fulfillment of the requirements
for the degree of Bachelor of Arts with Honors
“Sobald man über das bloße Konstatieren von Einzelheiten hinausgeht, sobald man versucht den Zusammenhang zu erfassen, die Erscheinungen zu begreifen, so betritt man auch den geschichtlichen Boden, wenn auch vielleicht ohne sich klar darüber zu sein.”
- Hermann Paul

“All the important changes have happened before the things they were supposed to change and it all sorts itself out in the end. The major problem is quite simply one of grammar…”
- Douglas Adams
# Table of Contents

Acknowledgments........................................................................................................iv
Abstract..............................................................................................................................v
Pronunciation Guide..........................................................................................................vi
1. Chapter 1 – Introduction.................................................................................................1
   1.1 Introduction to the problem.....................................................................................1
   1.2 Chain shifts.............................................................................................................2
2. Chapter 2 – Theoretical explanations of push chains......................................................6
   2.1 André Martinet: functionalism in sound change ....................................................6
   2.2 Functionalism & Structuralism................................................................................9
   2.3 Gradualness of sound change................................................................................12
   2.4 Functional load......................................................................................................14
   2.5 William Labov: a non-functionalist model of sound change ..................................17
   2.6 Problems with Labov’s account .........................................................................20
   2.7 Juliette Blevins: Evolutionary Phonology..............................................................24
   2.8 Synchronic chain shifts and modern functionalism..............................................27
   2.9 Interim conclusions..............................................................................................29
3. Chapter 3 – Greek & Portuguese....................................................................................31
   3.1 Prehistory: the Attic-Ionic vowel system.............................................................31
   3.2 Evidence & chronology.........................................................................................35
   3.3 A parallel situation: São Miguel Portuguese.........................................................40
   3.4 Analysis..................................................................................................................43
4. Chapter 4 – Swedish & Norwegian................................................................................49
   4.1 Prehistory: the Old Scandinavian vowel system..................................................49
   4.2 Evidence & chronology.........................................................................................51
   4.3 Analysis..................................................................................................................55
   4.4 Interim conclusions: categorizing Pattern 3 chain shifts.....................................56
5. Chapter 5 – The Northern Cities....................................................................................60
   5.1 Methodology.........................................................................................................60
   5.2 Completed phases: Labov et al..............................................................................62
   5.3 Ongoing phases: Eckert........................................................................................65
   5.4 Analysis..................................................................................................................67
6. Chapter 6 – General discussion & conclusions..............................................................71
   6.1 Generalizations about attested “push chains”......................................................71
   6.2 How do chain shifts begin?..................................................................................73
   6.3 What keeps chain shifts going?.............................................................................74
   6.4 Looking to the future............................................................................................77
7. References......................................................................................................................79
Acknowledgments

First, I am extremely grateful to the professors who helped me with this thesis over the past year. I very much appreciate how generous they have been with their time and their wise counsel, and I thank them for indulging my desire to pursue a project that is unusual by both Indo-Europeanists’ and theoreticians’ standards. Thanks in particular to Jeremy Rau, my advisor through it all, for commenting on innumerable drafts and for being a constant source of much-needed optimism and cheer even when progress was slow and painful. I am also indebted to Jay Jasanoff and Cedric Boeckx, who were of great help at several key points in the process. Thanks to Jay for helping with the Scandinavian chapter especially, and for sharing many entertaining anecdotes with me. Thanks to Cedric for guidance and support at every step, and for being an inspiration to me professionally.

Several administrators, both in the linguistics department and Dudley House, have also been extremely kind to me. Thanks to Cheryl Murphy and Allison Dimond for making me feel at home on the third floor, and for generally being wonderful, helpful people. Thanks to Tom Dingman, Laura Johnson, Orquidea Martinez, and Carvina Williams, the Dudley House team past and present, for their encouragement and assistance.

I would also like to thank the numerous professors who have helped shape my views about chain shifts and about linguistics in general: Adam Albright, Juliette Blevins, Andrea Calabrese, Edward Flemming, Andrew Garrett, Morris Halle, Norbert Hornstein, Jim Huang, Bill Idsardi, Michael Kenstowicz, Craig Melchert, Andrew Nevins, Colin Phillips, Oktor Skjaervo, Donca Steriade, and Bert Vaux. Thanks also to Stephen Mitchell, Christer Platzaack, and Tomas Riad for help with Scandinavian references.

My colleagues at Harvard, both undergraduates and graduate students in the department, have filled these four years with friendship and fun. I have benefited greatly from the vibrant intellectual climate in Boylston Hall, and I am glad to have many fond memories of class discussions, frantic study sessions, and LinG parties. On that note, thank you to the non-linguistic friends (!) with whom I’ve bonded over numerous animated debates, river walks, late-night jam sessions, Mario Kart marathons, and fine dessert liqueurs.

I owe special thanks and apologies to Stephen (and not just for correcting an embarrassing error in this sentence). Ever since I’ve known him, one of us has been working frantically on a thesis. It’s time to put an end to that! Thank you, in short, for being the best companion I could possibly want. Now back to life as usual!

Last and definitely not least, I owe a huge debt of gratitude to my family, especially my parents and the Hawkridge crew (who, incidentally, make excellent copy editors). Thank you for giving me a love of learning—and of languages!—and, quite simply, for making everything in my life possible. Thank you for not disowning me when I decided to take up “hysterical phrenology” instead of medicine. Thank you, Mom, for coming up with a catchy, Kremlin-on-the-Charles-y title. Thank you all for being so incredibly generous with your support, both financial and emotional. I cannot thank you enough.
Abstract

This thesis is an investigation of the controversy surrounding vocalic push chains, first proposed in phonological literature by André Martinet in 1952. Chapter 1 provides a general introduction to the push chain phenomenon, situating the development of chain shift theory within the larger picture of mid-twentieth century historical linguistics. In Chapter 2, I give an overview of the expectations of different theories of historical phonology with respect to the typology of chain shifts and predictions about push chains in particular. This involves looking at broader debates in diachronic phonology about the permissibility of teleology in sound change, the proper role of functionalism in phonology, the influence of lexical load, and the nature of sound change. Chapters 3-5 contain discussion of the four chain shifts that are commonly alleged to be push chains—Attic-Ionic Greek, São Miguel Portuguese, Old Scandanavian, and the American Northern Cities. I re-analyze each of the chain shifts, providing evidence that these cases do not in actuality constitute push chains, and I argue from this empirical evidence that push chains do not exist. In Chapter 6 I develop this conclusion more fully, explore the ramifications of these findings for phonological theory, and suggest directions for future research.
This thesis deals with precise classifications of speech sounds, so the use of technical notation which may be unfamiliar to the non-specialist is necessary. Here is a pronunciation guide for all the symbols that occur in the text. In most cases, I follow the International Phonetic Alphabet, except where convention differs for a particular language (e.g., Greek). Some materials I reference were written in non-standard notation, and in those cases I have preserved the original transcription system of the author. These variants, too, are noted here. I include a brief introduction to the articulatory and acoustic classification of vowels in the text itself.

[ ] indicates a precise phonetic transcription
// // indicates a phoneme, or contrastive speech sound. Slashes may indicate a higher level of abstraction than brackets, though the two may also be equivalent.
<> indicates a spelling in the script used by the language in question
: after a phonetic symbol or a macron (¯) over it denotes a contrastively long vowel
* before a phonetic symbol indicates a reconstructed (unattested but assumed) sound

[i] indicates the vowel in keep without diphthongization (high, front, unrounded, tense)
Labov’s /ih/ is a diphthong with nucleus [i] and a centralized offglide (i.e., [ia])
[u] indicates the vowel in boot (high, back, rounded, tense)
[ü] indicates the vowel in French lune (high, front, rounded, tense)
[i] indicates the vowel in sit (high, front, unrounded, lax)
Labov’s /i/ is equivalent to [i]
[e] indicates the vowel in day without diphthongization (close mid, front, unrounded, tense)
[e] is equivalent to [e]
[o] indicates the vowel in French peu (close mid, front, rounded, tense)
[e] indicates the vowel in get (open mid, front, unrounded, lax)
[e] and Labov’s /e/ are equivalent to [e]
[o] indicates the vowel in go without diphthongization (close mid, back, rounded, tense)
[ơ] is equivalent to [o]
[œ] indicates the vowel in score (open mid, back, rounded, lax)
[œ] and Labov’s /oh/ are equivalent to [œ]
[æ] indicates the vowel in cat (low, front, unrounded, lax)
Labov uses /eh/ to indicate a tense variant of [æ] found in some English dialects
[u] indicates the vowel in book (high, back, unrounded, lax)
Labov’s /u/ is equivalent to [u]
[A] indicates the vowel in but (open mid, back, unrounded, lax)
[a] indicates the vowel in lock (low, back, unrounded, tense)
[a] here may be taken as equivalent to [a], though in IPA it is low, central, unrounded
Labov’s /o/ is equivalent to [a]
[j] indicates the sound usually spelled <y> as in yak
[w] indicates the sound usually spelled <w> as in win
[j] indicates the diphthong in tree, with a nucleus [i] and offglide [j]
[ej] indicates the diphthong in way, with a nucleus [e] and offglide [j]
[ow] indicates the diphthong in snow, with a nucleus [o] and offglide [w]
[ew] indicates a diphthong not found in English, with a nucleus [e] and offglide [w]
1. Chapter 1 – Introduction

1.1 Introduction to the problem

In a monograph entitled *Sincronía, Diachronía e Historia: El Problema del Cambio Lingüístico* (1958), the Romanian linguist Eugenio Coseriu identified three problems historical linguists face: the “rational” problem of why languages must change, the “general” problem of identifying the conditions under which certain changes occur, and the “historical” problem of explaining changes that have already taken place. To this list, Weinreich, Labov, and Herzog (1968) add the “actuation” problem: why does a particular change occur in one language at a given time, but not in other languages with the same feature, or in the same language at a different time? Roger Lass (1976) calls attention to yet another question, the “inception” problem of determining the initial impetus for a given linguistic change.

These five problems constitute one of the main avenues of inquiry in historical linguistics. Uncovering the factors and mechanisms behind them has proven difficult—particularly so in the field of historical phonology. The aim of this thesis is to advance understanding of one vexing phenomenon in historical phonology, the push chain, with reference to the rational, general, historical, actuation, and inception problems as they pertain to this type of phonological change. As we will see, push chains are interesting from the perspectives of all of these problems: there are no circumstances under which they necessarily occur, the circumstances under which they may occur are far from clear, the mechanism by which they occur has not been explained, and the reasons why they occur seem murky. I will argue that the reason for all this mystery surrounding push chains is very simple: they simply do not exist, at least not as commonly understood.

In our quest to solve the mystery, we will evaluate the empirical and theoretical evidence for and against push chains. After a bit of background, namely a brief typology of chain shifts, we will turn to the question of why there are so few known examples of
push chains. We will sift through the claims of various theories with respect to the
distribution of push chains and discover why the theoretical claims underlying the
arguments for push chains are fundamentally unsound. I will then undertake analyses of
the four most commonly cited examples—Attic-Ionic Greek, São Miguel Portuguese, Old
Scandinavian, and the American Northern Cities—and explain the developments of these
systems in a way that does not appeal to a push chain mechanism.

I hope to show that these few cited examples of push chains do not constitute
sufficient evidence on which to make serious theoretical claims, and encourage my
readers to conclude along with me that there is no reason to posit a special push chain
category on either theoretical or empirical grounds—in fact, there are several good
reasons why not to posit one.

1.2 Chain shifts

André Martinet (1952; 1955), building on foundations laid by the
Neogrammarians and later Structuralists, introduced the concept of the phonemic chain
shift into the common linguistic vocabulary in the 1950s. The observation Martinet
brought to his readership’s attention is very simple: sometimes one sound change triggers
another, and this secondary change can trigger yet another. Lyle Campbell characterizes
this phenomenon very broadly:

“Sometimes several sound changes seem to be interrelated, with more far-
reaching impact on the overall phonological system of the language. These
changes do not happen in isolation from one another, but appear to be
connected, dependent upon one another in some way. Such interconnected
changes are called chain shifts. Several reasons have been put forward
why chain shifts should occur, and the final word about this is surely yet
to come…. ” (Campbell, 1999: 44)¹

¹ As Jay Jasanoff has pointed out to me, this comes close to defining the notion of
“phonological conspiracy,” which has been a topic of much discussion in recent years.
In practice, the characterization of chain shifts used by most linguists since Martinet’s time is narrower. A phoneme necessarily moves towards or away from other phonemes when it shifts, and in either one of these situations, the shift can trigger further changes. The phonological system of the language may change as a result, leading to new allophonic or phonemic relationships, or the system may retain a one-to-one mapping of pre-shift phoneme to post-shift phoneme, with only the phonetic realizations of those phonemes and the organization of the system changing.

Crucially, in order to be a true chain shift, there must be a causal relationship between each stage of the shift and the stage that follows it. As William Labov notes, “though the definition of chain shifting does not establish a direction of causation, the linguistic context in which these events occur establishes a clear causal relation” (1994: 119). This seems too optimistic a view—clear causation is often lacking in chain shifts, as we will soon see in our three case studies. At the opposite end of the spectrum from Labov, holding views far more conservative than mine on this issue, are those who subscribe to the view expressed by David Mortensen:

“There is little reason to believe that the historical processes that give rise to chains… are related at all. Two unrelated sound changes, the second of which coincidentally recreates some structure which had been obliterated by the first, can very easily give rise to scenarios of this sort. Furthermore, such chains do not show obvious progress in any dimension.” (Mortensen, 2004: 7)

In essence, the disagreement regarding causality stems from the actuation problem: there are no conditions that can be identified as being either necessary or sufficient to trigger a chain shift. This problem will figure largely in later sections, particularly in Chapter 2.

Let us now look at more concrete examples of what may happen when a phoneme, which we will call X, begins to shift. When X shifts into unoccupied phonetic
space, neighboring phoneme Y shifts into the position formerly occupied by the first phoneme. This is called a *pull* or *drag* chain, because the first phoneme seems to “pull” the others into the newly vacated space. In the other case, when X moves closer to Y, the two phonemes often merge into one, with the former distinction between X and Y being lost. There is also another possibility in Martinet’s theory: X’s movement towards Y may start a *push* chain, in which Y moves away from X, effectively escaping the impending merger. I schematize these three scenarios in (1.2.1) below. Additionally, theories which permit push chains also identify *mixed* chain shifts, with both push and drag components.

(1.2.1) Basic chain shift typology

i) A pull chain. The high vowels /i/ and /u/ diphthongize to /aj/ and /aw/ and then the mid vowels /e/ and /o/ raise to fill the vacated spaces. This is a simplified version of one part of the Great English Vowel Shift.

\[ 
\begin{align*}
& (1) \quad i \quad (2) \quad u \\
& \quad aj \quad e \quad a \quad aw
\end{align*}
\]

---

2 For our purposes, the relevant phonetic space is the area available to vowels. Though “experimental evidence supports the view that the [phonetic] space is perceptually defined in a way that does not precisely match either acoustic or articulatory measurements,” it can still be described in such terms (Stockwell and Minkova, 1988: 358). The vowel space is articulatorily bounded by the possible movements of the tongue tip and body. Tongue movement occurs across two axes, the vertical (height) and horizontal (backness), and is more easily controlled in the front of the mouth thanks to the tongue’s nimble tip. Height finds an acoustic correlate in a spectral energy peak called the first formant (F1). The lower the F1, the higher the tongue is in relation to the palate. Moving the tongue body back in the mouth lowers the second formant (F2), as does rounding the lips. The ear’s ability to distinguish differences in these acoustic parameters provides a perceptual limiting factor.
ii) A merger. The low vowel /a/ raises and fronts, first becoming /æ/ and then losing its distinctiveness as it continues on its trajectory, completely merging with /e/. This occurred in the long vowel system of Attic and Ionic Greek during the first millennium BCE.

iii) A hypothetical push chain. The low vowel /a/ raises, causing the mid back vowel /o/ to raise. The high back vowel /u/, under pressure from the encroaching /o/ but unable to raise any further, fronts to /ü/.

The pull or drag type of chain shift is widely attested in vowel, consonant, and tone systems. The other type, the push chain, has only been seen with vowels and is so rare as to cause alarm; linguists have long debated whether it actually exists. Some linguists have difficulty finding reasons why push chains should be theoretically expected yet defer to empirical evidence that seemingly supports push chains (Campbell and Ringen, 1981), others find the explanations for push chains too reliant on teleology and categorically deny the existence of push chains as a result (King, 1967, 1969b, a), and still others have no problem with the theoretical underpinnings of the concept but lament the lack of attested examples (Hock, 1986; Trask, 1996). Recently, Labov (1994 et seq.) has also attempted to explain push chains without appealing to teleology. In the discussion to follow, we will evaluate all these positions as well as the examples that linguists have cited over the past fifty years as empirical evidence in favor of the existence of push chains.
2. Chapter 2 – Theoretical explanations of push chains

2.1 André Martinet: functionalism in sound change

In a 1952 *Word* article called “Function, Structure, and Sound Change” and the subsequent book *Économie des Changements Phonétiques*, André Martinet introduced the idea that causally linked phonemic drifts come in two flavors, the now-familiar push (“chaîne de propulsion”) and pull (“chaîne de traction”) varieties. Writing at a time when efforts to monitor sound change in progress were nonexistent, Martinet cautiously noted that it can be difficult to determine the exact nature of a chain shift’s chronology (and, therefore, its classification as push or pull) retrospectively because such structure-preserving “rephonologization” changes³ tend to leave behind only scant orthographic evidence.

Martinet also observed that push shifts differ from pull chains in one crucial theoretical respect: the assumption that one phoneme can move away from an encroaching phoneme to preserve a margin of security brings an element of teleology into the theory. In other words, functional considerations, namely the desire to retain comprehensible speech by avoiding confusing mergers, seem necessary to explain push chains. “This type of assumption conflicts of course with the traditional views concerning the ‘blindness’ of ‘phonetic laws,’” Martinet wrote, but he was content to allow the teleological explanation because he found that it was “not too difficult to understand how a phoneme can yield under the pressure of one of its neighbors” (1952: 132)⁴. As a general principle, this statement lies at the very heart of virtually every theory of push

---

³ In Jakobsonian terminology, structure-preserving changes to a language’s phonemic system are categorized as “rephonologization,” as opposed to secondary phonemic split (“phonologization”) and merger (“dephonologization”).

⁴ All page numbers for “Function, Structure, and Sound Change” are from the reprinted edition in Baldi & Werth (1978).
chains: a phoneme must somehow be able to look ahead and determine that it must move in order to avoid merger.

The concept of the chain shift ties in well with Martinet’s conception of linguistic economy. Building on the work of Henry Sweet (1874), Martinet viewed language change as the result of tension between “communicative needs” (i.e., “the preservation of useful phonemic opposition”) and the inherent laziness of the speaker, which would turn language into unintelligible mush if left unchecked (Martinet, 1952: 126). These principles underlie catch phrases commonly used in both diachronic and synchronic phonology today, such as “maximization of perceptual contrast” and “minimization of articulator effort.” One theory that relies heavily on contrast maximization is Dispersion Theory, which purports to explain how phonemic systems organize themselves for perceptual ease. Because Dispersion Theory is more comprehensible after one has a grasp of concrete patterns of sound change, I will refrain from discussing it until Chapter 6.

Another force that is often viewed as a complement to dispersion is functional load, which prevents mergers that would create rampant homophony. Functional load is the subject of §2.4.

In his *Word* paper, Martinet provides only one example of a push chain: [u] fronting to [ü] followed by [o] raising to [u] and other subsequent changes in the Azorean Portuguese dialect of São Miguel. Despite using this case as his illustration of a push chain, Martinet admits that even this case shows how blurry the distinction between push and pull chains can be (1952: 133). As Anttila notes, “some consolation is given by the fact that both chains represent one and the same underlying force, even if this is controversial and not perfectly established yet” (1989: 186). The pressure which, in Martinet’s popularized view, contributes to the initial movement of [u] to [ü]—that is, the “push” part of the chain—is not any actual encroachment by a newly-moving phoneme but rather the mere existence of “too many” back vowels. In this case, the language had four back vowels altogether before the shift. The rest of the chain, with former /ø/, /œ/,
and /a/ raising in sequence, clearly involved the pull mechanism. I argue in Chapters 3 & 4 that virtually the same sequence of events also occurred in Greek and many other languages, but that most of them (including the Greek case, in one instance) typically get classified as pull chains. William Labov and his colleagues (Labov et al., 1972; Labov, 1994: 129) even gives the [u]-fronting chain shift a special name: the “Pattern 3 shift.” Understanding exactly what underlies these examples will prove critical for constraining our theory of what types of chain shifts may occur.

Though Martinet’s introduction of chain shifting was a large step forward for historical linguistics, it also raised many questions about the nature of sound change and how to reconcile chain shifting with the rest of diachronic phonology. I sketch some of these questions briefly below; we will return to each of them.

(2.1.1) Questions raised by Martinet’s theory of chain shifts

- Do functional/teleological principles guide phonology?
  - If so, is teleology really necessary to explain push chains?
- Can push chains work in featural models of sound change?
- Are push chains empirically attested?
  - If so, why are push chains so rare compared to drag chains?
  - Under what conditions do push chains, as opposed to mergers, occur?

Confronted with these questions, various linguists have come to very different conclusions. A survey of the half-century of literature that has accumulated since Martinet’s original proposal shows that virtually every approach has been explored. Some linguists have adopted Martinet’s theory wholesale, though often with worries about the scarcity of confirmed push chains. Others deny the theory and the empirical evidence. A couple more recent authors, such as William Labov and Juliette Blevins, accept the examples but seek to explain them without an appeal to teleology. I hope that after taking a closer look at the theoretical and empirical evidence, you will share my conclusion that both the Martinetian and non-teleological theories that have been proposed fail to explain push chains, but that the purported examples can easily be explained in a way that is properly constrained, virtually neutral towards both synchronic and diachronic
phonological theories, and reliant only on non-controversial, independently motivated principles.

2.2 Functionalism & Structuralism

To answer the first question posed in (2.1.1), whether one should be concerned by the teleological nature of Martinet’s explanation of push chains, a bit of history may be useful. Outside the realm of linguistics, teleology has come in and out of fashion at various times through the centuries: “for Aristotle, a non-teleological universe (even with respect to inanimate matter) was inconceivable… from Darwin on, teleology is pretty much anathema, or at the very least weak-minded, romantic or obscurantist” (Lass, 1980: 64). This is certainly not the case in linguistics today, nor was it the case in the mid-twentieth century. During Martinet’s heyday, this issue was quite contentious. The Prague School linguists, based primarily in Eastern Europe (with a few notable exceptions, including Martinet) during the first half of the twentieth century, were particularly amenable to teleology in diachronic phonology, since in their view all of language necessitated consideration from a functional perspective. Jan Baudouin de Courtenay (1845-1929) and Otto Jespersen (1860-1943) can be viewed as the intellectual forefathers of the functionalist movement, though it did not gain momentum until a couple of decades into the twentieth century. Roman Jakobson was one of the first linguists to develop the fledgling theory more fully. Twenty-five years before Martinet introduced the push chain, Jakobson opined in a Prague Linguistic Circle paper that “the overlapping between territorially, socially or functionally distinct linguistic patterns can be fully comprehended only from a teleological point of view, since every transition from one system to another necessarily bears a linguistic function” (1962 [1927]: 1). He also later wrote that “quand nous considérons une mutation linguistique dans le contexte de la synchronie linguistique, nous l’introduisons dans la sphere des problèmes téléologiques” (1962 [1931]: 218).
Such acceptance of teleological explanation was by no means the consensus outside of the Prague School, however. In other circles, the approach was met with strong criticism:

“While Jakobson’s propositions diverged from the practice of other linguists in all of the major respects, this was especially true in his urging a concentration on the system of distinctive sound differences to the exclusion of other phonetic facts, and in proposing a teleological, system-determined conception of linguistic change. It is by no means clear that the latter notion ever really prevailed: while historical studies came soon to be cast in terms of changes undergone by the phonological system, the role played by the system in motivating change generally in a teleological fashion was stressed more by theoreticians (e.g. Martinet) than by the mainstream of practicing historical linguists (which is not to deny that Martinet himself did substantive work of a historical nature).” (Anderson, 1985: 89)

Saussure sought to maintain a strict separation between synchrony and diachrony, a dichotomy which Jakobson rejected because it precluded the possibility of interpreting linguistic change teleologically (Anderson, 1985: 118). And Leonard Bloomfield, in stark contrast to Jakobson and Martinet, called teleology “a mentalistic pseudo-solution” that “cuts off investigation by providing a ready-made answer to any question we may ask” (1970 [1934]: 284).

The debate over teleology in linguistics has been particularly contentious because it ties into a deeper question, namely whether formalism or functionalism is the correct approach. This issue has occupied several generations of linguists, at least as far back as Edward Sapir (1884-1939) and Nikolai Trubetzkoy (1890-1938). As it happens, functionalism fell out of favor during the mid-twentieth century with the advent of generative phonology, but in the early 1990’s it came back with a vengeance. Today the

5 Linguists are often divided into competing “functionalist” and “formalist” camps, though there is hardly consensus about what these labels actually mean. For our purposes, it should suffice to say that formalists typically believe that language is characterized by autonomous organizing principles not shared by other cognitive systems, while “functionalist linguistics is based on the assumption that language is the way it is because of the functions to which language is put” (Nathan, 1999: 309).
predominant theory of synchronic phonology, Optimality Theory (Prince and Smolensky, 1993), makes extensive use of functionalist principles.

Some linguists who have a theoretical stake in the functionalism-formalism debate have unfortunately allowed their preconceived notions about teleology to cloud their vision. Those who seek to eliminate teleology, when confronted with the assertion that push chains require a teleological explanation, by and large move to deny the empirical data. If there are no push chains, then there is no problem. This is the tack taken by King (most explicitly in 1969b), McGinn (2005), and Vincent (1978), to name an outspoken few. On the other hand there are those who believe that the issue is purely empirical and who trust that the data supports the existence of push chains. Advocates of this approach include Anttila (1989), Campbell & Ringen, who state that “the issue is not teleology, but rather the empirical validity” (1981: 65), Trask (1996), and Labov (1994), whose proposal we will examine in detail later in this chapter.

The goal of this paper is not to advance any particular theoretical agenda; however, at this juncture it is difficult to avoid doing so. My philosophy is that the relationship between empirical evidence and theory should be reciprocal. Having too many theoretical encumbrances prevents objectivity when assessing data, but pure empirical descriptivism without theory leaves much to be desired when it comes to utility. I am well aware of the inherent trade-off between flexibility and explanatory power in a theory—the more specific a theory’s predictions, the more it can explain, but it is also easier to prove such a theory wrong. This is sometimes called “the constraints problem” (Labov et al., 1972). My solution to this problem is to posit only what is required to account for the data and avoid making predictions that overgenerate, while at the same time realizing that not all possible forms are necessarily attested. By avoiding as much theoretical commitment as possible in this way, it is still possible to salvage something if new empirical evidence comes along or as the fashion in phonological theory changes. Thus, I have set out to formulate a picture of push chains that dispenses with as many theory-internal demands as possible.
There is nothing wrong with teleological explanations *a priori*, so the deciding factor must be evidence from real languages. Strictly on grounds of parsimony, it is desirable for phonology to do without teleology, just as it is desirable for phonology and syntax to do without derivational look-ahead, but language can be surprising. If there is incontrovertible evidence that push chains exist, and that they cannot be explained except through teleology, then so be it. In Chapters 3, 4, and 5, we will have a look at the cases that have been co-opted in support of the functionalist view so as to arrive at an informed conclusion about teleology.

2.3 Gradualness of sound change

Another theory-internal reason for unease regarding push chains is the lingering belief that phonological change is discrete. The introduction of a full-fledged featural theory\(^6\) in *The Sound Pattern of English* (Chomsky and Halle, 1968) created a very visible revolution in synchronic phonology, and it affected diachronic phonology in a way that was perhaps less noticeable but no less powerful. With the notion of features came the idea that sound change could be explained as featural change; this theory correctly predicted the ubiquity of sound changes involving the switch of a single binary featural value. Under a strict interpretation of featural theory, discrete featural change is the only admissible type of sound change—this falls directly out of the idea that what we call a “phoneme” is just a shortcut for specifying what is represented in the mind as a feature bundle. This seemed like a very good idea to a lot of historical linguists who were already favorably disposed towards Structuralism, since “the distinctive feature

---

\(^6\) In featural theory, speech sounds are mentally represented not as atomic units but rather as bundles of simultaneously occurring, distinctive features specifying various aspects of their articulation. For the present purposes, all features can be assumed to have binary values, which are represented with + and −, indicating the presence or absence of the particular feature in question. The relevant features for our discussion of vowel shifts are \([\pm \text{high}], [\pm \text{low}], [\pm \text{back}],\) and \([\pm \text{round}],\) which specify tongue height, tongue backness, and lip rounding.
framework makes it possible to state precisely and economically the effect of a change on the membership and/or internal structure of the system as a whole” (Bynon, 1983: 81).

The featural theory of sound change cannot admit push chains because if sound change is discrete, “it is not conceivable that one phoneme can encroach on the phonetic space of another (or in Martinet’s terms, reduce its margin of security)” (Labov, 1994: 200) without merging with it. Take, for example, a case of /o/ “encroaching” on an adjacent phoneme, /u/. The featural specification of /o/ is [+back], [-low], [-high], [+round]. The only way in which /u/ differs from /o/ is its specification for [+high]. Therefore, if only featural change is allowed, the raising of /o/ (i.e., a change from [-high] to [+high]) as the first step in a chain will necessarily cause it to merge immediately and irreversibly with /u/. This caused some concern about the plausibility of the push chain scenario and led some phonologists to reject it (cf. King, 1969b, a).

Following my policy of remaining neutral with respect to theory, I withhold judgment on the discreteness of sound change issue. This may seem strange to readers who follow phonological theory and know that the gradualness of sound change is now commonly assumed. Much has changed since Robert King’s proclamation that “as fondly held as the gradualness assumption about sound change seems to be, there is hardly a shred of empirical support for this claim” (King, 1969b: 6). Many phonologists today think of sound change in terms of gradual movements by exemplar-based categories, not binary switching of feature values (cf. Labov, 1994; Aitchison, 2001; Pierrehumbert, 2001; Blevins, 2004). However, as we will see, this debate cannot be decided by evidence from completed chain shifts. None of the documented examples of push chains actually contain “encroachment” of the type that would prove problematic for the featural

To be fair to the featural theory, we should note that the existence of synchronic chain shifts (see §2.8) shows that chain shifting can happen instantaneously as the result of feature manipulation, which casts doubt on the need to postulate the gradualness of sound change in order to account for diachronic chain shifting. See Idsardi & Son (2004) for one such synchronic chain shift case in Korean.
theory of sound change. This fact will prove critical to the new analysis of push chains presented here, and discussion of the issue will resume in Chapter 6.

2.4 Functional load

Martinet realized that merely identifying push chains is insufficient; one must also provide an explanation for the complementary distribution of mergers and push chains. This is easier said than done, however. The idea of homophony avoidance was first introduced by Jules Gilliéron (1918), and the principle was soon adopted and expanded by various other European linguists. Borrowing an idea from Vilém Mathesius (1929; 1931) that had been popularized in the Prague Circle by Trnka (1931), Jakobson (1962 [1931]), and Trubetzkoj (1939), Martinet invoked a direct descendant of Gilliéron’s theory now known as the concept of functional load or yield. Functional load is based on the notion that “since phonemes, by definition, serve to distinguish between words and forms, any phonemic merger will inescapably involve confusions detrimental to the normal functioning of the language,” so in cases where that confusion would be too detrimental to the language, the theory predicts that a push chain would occur instead of a merger (Martinet, 1952: 128-129).

Since the introduction of the lexical load hypothesis, linguists have made two types of objections to this type of explanation. Some question the explanatory power of the principle:

“The homonymy-prevention theory contributes little to the solution of the ‘actuation riddle.’... it is not clear that a theory based upon the functional yield of cognitive contrasts can provide the machinery for assessing the full complexity of causal relations within phonological structure.”
(Weinreich et al., 1968: 137)

More specifically, many phonologists worry that the ubiquity of phonemic mergers, even when rampant homophony results (just listen to French if you disagree), casts doubt on the validity of any principle purporting to prevent homophony (cf. King, 1967, 1969b, a;
Bynon, 1983). If language abhors merger, then why are mergers so common and push chains so rare? This is a very difficult question to answer unless one retracts the assumption and grants that mergers are not in fact avoided whenever possible.

Even those who accept the lexical load hypothesis in principle still qualify their endorsement with one big caveat: nobody knows what lexical load is. In his survey of the subject, Meyerstein provides no less than ten different definitions from the forty years spanning between 1930 and 1970, each proposed by a different author (1970: 15-17). Perhaps even more crucially, there are many possible ways in which functional load might be quantified. “For lack of yardsticks the notion of functional load was of questionable descriptive relevance” during Martinet’s time and for some years afterwards; even now, we are faced with an overwhelming number of possibilities to calculate, and no clear way to determine which is correct (Meyerstein, 1970: 26). Bynon divides this problem into two parts:

“Firstly, there is the question of the linguistic context within which the functional load of a contrast either was lost, or would have been lost, as the result of a particular merger is to be measured. Should it be phonological environment in running texts? or minimally distinct word pairs in the lexicon? Secondly, should functional load be quantified in terms of phonemes, (allo-)phones or features?” (1983: 87)

A number of linguists invented and tested mathematical algorithms for quantifying functional load during the 1950s and 1960s. The group of investigators included Joseph Greenberg (1959) Charles Hockett (1955; 1967), Robert King (1967), Henry Kučera (1963), R.S. Meyerstein (1970), Jørgen Rischel (1961), and William Wang (1967). Despite the number of researchers attacking the problem, very little progress towards the goal of explaining mergers and chain shifts was actually made. Some of the studies explicitly denied that the results were applicable to diachronic phonology (i.e., Hockett, 1955, 1967), and others could not be extended beyond the particular languages for which they were designed (i.e., Kučera, 1963). Of all these studies, only King (1967)
focused on creating a statistic to handle the diachronic facts, and his results did not lend support for the theory that functional load plays an important role in potential mergers. Of course, given the fact that King merely studied a small subset of the many possible formulations one could conceive, it would be foolhardy to treat those results as definitive.

Much more recently, Dinoj Surendran and Partha Niyogi (2003; forthcoming) have formulated more sophisticated algorithms with the help of computational power not available in earlier decades. Using the merger of /n/ and /l/ word-initially in Cantonese as their case study, Surendran and Niyogi found that the functional load of the opposition in question cannot be the sole diagnostic for predicting merger. Out of 171 consonantal oppositions in Cantonese, 74% had lower functional load than the n-l opposition, the one that ultimately merged. The authors note, however, that some of these oppositions do not seem relevant (they fail to elaborate on this point) and propose that “only those pairs that are likely to merge should be considered” in the comparison of functional load statistics, and that “likely to merge” could ostensibly be defined as sharing a common place of articulation (forthcoming: 3). This brings us back to where we started: functional load statistics cannot be interpreted in the abstract, because it is still unclear what “high” and “low” functional load are. Also, in Surendran and Niyogi’s one test case, the merged opposition did not seem to have low functional load as compared to the rest of the oppositions in the language, as Martinet and his followers would predict.

The inconclusiveness of mathematical investigations of functional load is discouraging, but I do believe it is worthwhile to think carefully about what one would hope to gain from such studies if they were to come out more in line with functionalist expectations. After providing a series of quotations from authors who stake diachronic claims—such as the preference for chain shift over merger—on functional load, Meyerstein makes a very salient observation:
“Functional load may be a factor in sound change, but there are other factors which may reduce or, in fact, undo the effect of functional load. This can scarcely be called an unequivocal claim, and as ‘iffy’ an effect as that suggested [by proponents of functional load]... is hardly amenable to quantificational checks which recent scholarship has demanded and, indeed, attempted.” (1970: 20).

This is quite a strong condemnation of the enterprise, coming from the first chapter of a book on the quantification of functional load! It is certainly worth keeping in mind that, even if it were possible to ascertain the precise nature of functional load as calculated mentally by a language learner (if it is indeed true that a language learner does such a thing), there would still be much left to desire. Unless and until researchers can narrow down the possibilities and develop a formula with some predictive value, neither of which I consider very likely, there is no information from functional load that helps answer questions about push chains.

Now that we have discussed all the basic Structuralist principles underlying Martinet’s push chain mechanism, we are ready to leave the Prague School and move forward in time. In sections 2.5-2.8, we will survey some theories that have been developed over the past decade. They provide an interesting contrast to the material presented in this chapter; while more recent theories abandon some Structuralist tenets in favor of phonetic and perceptual principles, they also appeal to some time-tested principles that by now should seem very familiar.

2.5 William Labov: a non-functionalist model of sound change

The first “new” theory of sound change we will discuss has actually been evolving for several decades. The larger program of which this is part dates back to articles by William Labov in the 1970s, but it is most fully developed in Labov’s magnum opus, *Principles of Linguistic Change (PLC)*, a projected three-volume set of
which the first installment appeared in 1994. The PLC model appears attractive because it takes sociological factors into account and makes use of quantitative empirical data from chain shifts that are currently in progress. In the PLC volume on the internal factors of linguistic change, Labov defines a ‘minimal’ (i.e., two phoneme) chain shift as “a change in the position of two phonemes in which one moves away from an original position that is then occupied by the other” (1994: 118). Labov explicitly rejects Martinet’s (1952) teleological explanation of chain shifts while still allowing push chains. He also makes explicit three principles of vowel shifting, which he claims can account for most vocalic chain shifts. These are listed in (2.5.1) below.

(2.5.1) Principles of vowel shifting (Labov, 1994: 116)

PRINCIPLE I – In chain shifts, long vowels rise.
PRINCIPLE II – In chain shifts, short vowels fall.
PRINCIPLE IIA – In chain shifts, the nuclei of upgliding diphthongs fall.
PRINCIPLE III – In chain shifts, back vowels move to the front.

In PLC I, Labov explicates his model of sound change, which he proposes as an alternative to the functionalist account. Labov’s scheme builds on the argument proposed by Anthony Kroch (1989a; 1989b) for syntax, that misunderstandings drive linguistic change. The model of sound change put forward in PLC utilizes misunderstandings to propagate shifts in an exemplar-based model of perception (cf. Pierrehumbert, 2001 for a more detailed overview of one rendition). In an exemplar-based model, each phoneme or allophone corresponds to a category of remembered tokens, with each token encoding both acoustic and non-acoustic data (e.g., the surrounding phonemic environment, the speaker’s gender and social position, etc.) about one particular instance of that phoneme’s occurrence. Each token contributes to statistics about the category, which are constantly

---

8 The first volume, to which we will make extensive reference, deals with the internal factors of linguistic change. The second volume, published in 2001, focuses on the social factors. A forthcoming third volume, expected in 2006, will treat the cognitive factors.
updated during the phonological acquisition process, with more recent and/or salient tokens carrying the most weight. The mean acoustic values of all the tokens in a category then becomes the target for production of that category.

To understand how Labov uses an exemplar-based theory to model sound change, take, for example, the word block. If a listener hears an aberrant token of the word block containing an /o/ which falls in the phonetic space between /o/ and /æ/ (i.e., the token sounds somewhat like black), he may fail to comprehend the word and therefore not include the token’s parameters in his running calculations of the mean value of /o/. “If no sound change is involved, this mechanism of misunderstanding will have a conservative tendency, reinforcing the separation of /æ/ and /o/,” Labov claims (1994: 587). This scenario, with /æ/ and /o/ close together, is shown in figure (2.5.2) below. If /o/ is moving towards /æ/, discarding the aberrant tokens will retard this change.

(2.5.2)

The stable scenario, showing the mean F2 value of /o/, represented by the black square, as 1550Hz. The leftmost white square represents a misclassified aberrant token of /o/. Figure modified from Labov (1994: 586).

There is a second scenario, in which /æ/ and /o/ are separated by a larger margin because of a shift fronting and raising /æ/ to the vowel Labov loosely transcribes as /æh/. In such a case, an aberrant token of /o/ is less likely to be miscategorized as /æ/. The model here essentially formalizes Trask’s idea that “the farther apart the contrasting segments are in
phonological space, the easier it is to tell them apart when listening to speech, and the less likely are misunderstandings” (1996: 87). In this situation, an aberrant token, if identified correctly, will contribute towards the mean of /o/ and cause it to move in the direction of the already in-progress shift of /æ/. Thus, a chain shift is born—so the logic goes. This shift scenario is shown in figure (2.5.3) below.

(2.5.3)

The shift scenario, showing the mean F2 value of /o/, represented by the black square, as 1571Hz. The leftmost white square represents an aberrant but correctly classified token of /o/. Figure modified from Labov (1994: 586).

2.6 Problems with Labov’s account

Though Labov’s aim of reconciling the Neogrammarian regularity of sound change hypothesis with lexical diffusion and exemplar theory is quite admirable, his model does not actually make correct predictions about the nature of sound change.

The PLC model is not consistent with the cross-linguistic generalizations that Labov and others have made about when chain shifts occur. In the model Labov proposes, the only relevant difference between the “stable” and “shift” scenarios in (2.5.2) and (2.5.3) is the large space between categories in (2.5.3). There is, according to Labov, a higher probability of /o/ moving in (2.5.3) because of the distance of its
neighbor, regardless of whether that space is the result of a shift by /æ/ or not. Critically, this is not an explanation of pull chains at all, because the movement of /o/ is not causally linked to any prior or ongoing movement by /æ/. Furthermore, taking the model seriously, one would predict that systems with few vowels or relatively small, concentrated vowels would be less stable than systems with more crowded vowels. This prediction actually proves quite contrary to Labov’s statements elsewhere in PLC—following Martinet, he invokes crowding of the phonemic space as a major cause of instability and the force behind many chain shifts, in particular those which abide by Vowel Shifting Principle III (cf. Martinet, 1955; Labov, 1994: 118).

I also call attention to the lack of difference between the stable and shift scenarios, contrary to what Labov would like. The PLC model relies on the assumption that, upon hearing a very /æ/-like token of /o/ in the stable scenario, a learner misclassifies that token as /æ/. Such miscategorizations should have consequences—the learner might be expected to create a new lexical item with the same semantics as the misheard item but containing /æ/ rather than /o/, or posit some sort of vowel alternation between /æ/ and /o/. This is not what actually occurs. As Morris Halle (p.c.) rightly notes, anyone who has listened to someone with an accent (or received a letter from someone with poor handwriting) intuitively understands that context is typically sufficient to allow for decipherment of aberrant tokens; if decipherment is not possible, we observe no consequences other than lack of understanding. This being the case, one should not expect misclassification of outliers to impede change in the stable scenario. There should be no difference between the stable and shift scenarios with regards to the effect of outliers on the mean (and therefore, the pronunciation) of /o/.

Though Labov’s model of chain shifts eliminates the appeal to teleology for which Martinet receives such strong criticism, it does so at a very high cost. The only reason it avoids teleological explanation is that it does not explain push chains—and in fact, it cannot explain mergers, either! Stockwell and Minkova note that Labov’s definition of a minimal shift in the first volume of PLC only works for drag chains, but that this is
probably an unintentional oversight (1997: 285). In the second volume of PLC, Labov writes:

“Though [the model] explains the predominance of pull chains, it does not explain two facts: (1) not all holes in patterns are filled; not all chain shifts are activated at all times, and (2) push chains do occur (as when, for example, the backing of /e/ precedes the backing of /ʌ/ in the Northern Cities Shift). There must be a social force that activates the shift and drives the increment.” (2001: 463)

I argue that the push chains are not just unexplained by the PLC model, but actually prove to be incommensurable with it. If Labov’s explanation is correct, an incipient push chain will never be able to arise. The stationary phoneme would never have any reason to begin moving because the encroaching phoneme’s motion towards it would be retarded, as in the “stabilizing” scenario (2.5.2). This stabilization would also have the effect of preventing wholesale mergers from occurring because the vowels could never get close enough together to produce the level of confusion necessary for the loss of distinction. This is certainly not a desirable result, since mergers are quite well attested in the histories of the world’s languages.

Even though Labov’s model cannot actually admit push chains, he happily allows them. Labov succinctly states his position on the status of push chains in the first volume of PLC:

“In discussions of chain shifting, many abstract arguments have been put forward concerning the relations between the entering and leaving elements in a chain…. The theoretical issue that motivates the discussion has to do with the nature of sound change itself. Those who believe that language change can only arise through changes in discrete (perhaps binary) rules can accept only pull chains (King, 1969a). In this framework

---

9 In the same article, Stockwell and Minkova point out problems with Labov’s principles of chain shifting and his conception of “peripherality.” None of these issues bear on the push chain debate so I will not discuss them here, but I encourage the interested reader to consult Stockwell and Minkova’s paper for details.
it is not conceivable that one phoneme can encroach on the phonetic space of another (or in Martinet’s terms, reduce its margin of security). In the case of the backing of /e/, [-back] will either become [+back] or not become [+back]. But those who view sound change as taking place in a continuous phonetic and phonological space, as I do here, will see the issue of pull chains versus push chains as an empirical one…. Real-time evidence confirms these indications of a push chain status for the backing of /e/.” (1994: 199-200)

This overly simplistic assessment fails to capture the variety of differing views within the anti-push shift camp. While a strictly featural conception of sound change turns more structurally-oriented linguists away, there are plenty of other theoretical and empirical grounds on which people may find fault with the push chain scenario, as demonstrated earlier in this chapter.

Of course, Labov is right to note that those who take Structuralism seriously will be left unsatisfied by his approach. At the core of the PLC model is the idea that vowels somehow want to preserve the amount of area in the vowel space that they inhabit, preferring a geometric “translation” of that area to non-area-preserving deformations when circumstances warrant a change. Perhaps this explains why Labov focuses on chain shifting in vowels; it seems plausible that the vowel space is continuous, but the consonant space intuitively seems more quantized, and indeed there are no convincing examples of consonantal push chains (Bynon, 1983; Aitchison, 2001). Juliette Blevins shares this view, remarking that consonantal chain shifts of any type are rare because “consonant perception appears to be categorical and not structured with prototypes…. However, the phonoetically based classification of stops by phonetic features of VOT and burst amplification is potentially gradient, allowing for the same general treatment as the more common vocalic chain shifts” (2004: 291). Morris Halle, taking a different tack, rejects any notion of asymmetry between consonants and vowels—he criticizes Labov for making the vowel space seem “too psychologically real” and refuses to believe that chain shifting constitutes a real phenomenon distinguishable from random phonemic drift.
I disagree on both counts. Still, even if Labov’s idea about the psychological reality of the vowel space is true, it is difficult to see how a principle of area preservation could translate into a Hallean system of phonology that admits sound change only insofar as it is instantaneous grammar change begotten by the manipulation of features. Labov’s theory proves as poor a fit for feature-based frameworks as Martinet’s in this respect.

Labov’s model of sound change perhaps comes the closest to a concrete explanation of sound change of any non-feature based theory. Unfortunately, the theory falls apart under close examination, and it does not provide an explanation for causally linked sound changes—that is, chain shifts. Most crucially for the present investigation, though Labov is one of the most outspoken proponents of push chains, he has no theory to support his position. While other linguists cite Labov as having “solved” the lingering problem by explaining push chains in non-teleological terms (cf. Gordon, 2002: 253), he does not actually have a mechanism or theory capable of allowing push chains. Labov does, however, claim to add empirical support to the push chain notion by bringing another alleged case, the twentieth century American Northern Cities Shift, to the table. This shift, which I argue not to be a push chain, is the subject of Chapter 5.

### 2.7 Juliette Blevins: Evolutionary Phonology

Until very recently, Labov’s was the only real alternative to a Structuralist analysis of chain shifts. In 2005, Juliette Blevins provided another option, fusing phonetics with historical linguistics in her book, *Evolutionary Phonology*. The overarching goal of the book is to explain phonological phenomena without recapitulating historical changes in the synchronic grammar. Blevins primarily achieves this by eliminating appeals to principles of synchronic markedness, which form the very heart of Optimality Theoretic phonology. In some sense, Blevins’ theory of chain shifting resembles Labov’s. Both make use of exemplar theory to advance sound change, and both take the gradualness of sound change and the continuousness of the vowel space
very seriously. I quote below from a summary of the *Evolutionary Phonology* position on chain shifting:

“Vocalic chain shifts are the combined result of intrinsic variation with the prototype structure of vocalic categories. Chain shifts can arise naturally when a formerly occupied area of the psycho-acoustic space is opened up, with variation giving rise to better prototypes of a pre-existing category in the newly opened space.” (Blevins, 2004: 291)

This passage makes it clear that the *Evolutionary Phonology* model of chain shifts does not try to explain the first link in the chain, and indeed I will show in just a moment that it cannot explain one very common initial shift, [u] fronting to [y]. The model also fails to account for any shift in which one vowel moves anywhere other than the midway point between its neighbors, as must happen in the “encroachment” commonly thought to cause push chains. These limitations fall naturally out of the particular type of exemplar theory that Blevins adopts.

The basis of Blevins’ mechanism is Pierrehumbert’s (2001) exemplar-based model of perception, which *Evolutionary Phonology* uses to account for production facts as well. In this model, when a speaker wants to produce a vowel—[u], for the sake of argument—he attempts to produce the “best” exemplar of [u] that he has heard. Crucially, “best” in this context means most likely to be categorized as [u]. This statistic, Pierrehumbert claims, comes from the exemplar’s summed similarity to other exemplars of [u] taken as a fraction of its summed similarity to exemplars of all vowels.¹⁰ Now the other members of the vowel system become critical. As a test case, consider a five-vowel system of /i, e, a, o, u/. In this vowel space, [u]’s only close neighbor is [o]. There are no vowels higher or more back than [u], and [i] is distinguished from [u] by its lack of rounding in addition to its frontness. Because of the structure of the system, the best

¹⁰ One can certainly argue about the plausibility of keeping such mental statistics. I leave this matter to the reader’s judgment.
exemplars of [u] by Pierrehumbert’s rubric would in fact be closer to the edge of the vowel space than the mean exemplar of the [u] category, because these exemplars would be least confusable with [o]. Blevins translates this perceptual effect to production by stipulating that speakers try to produce the best exemplar of a given category. This feedback loop of perception and production has the effect of distributing vowels evenly throughout the perceptual space, because the system will reach equilibrium only when the best exemplar of each category is also its mean.11

Since [u] is closer in the perceptual space to [o] than it is to [i], the Pierrehumbert/Blevins model can indeed capture the mechanism by which /u/ would begin to front. One would in fact expect this fronting to happen in every vowel system in which [i] is not the nearest neighbor to [u]. This may explain why /u/ exhibits some degree of fronting in a wide variety of languages and dialects, a partial list of which is presented in Chapter 6. More problematic, and indeed impossible to explain using solely the “best exemplar” theory of sound change, is ascertaining why /u/ would front all the way to [y] rather than stopping at the point maximally distant from both [i] and [o].

The model runs into the same problem when explaining the intermediate steps of a push chain, because such shifts would require encroachment. In other words, because vowels must always move away from one another in this model, it cannot account for situations in which /o:/ raises (almost) to [u:] prior to the fronting of /u:/ to [y:], but it can account for situations in which /u:/ fronts and then /o:/ later raises. Blevins unequivocally states that “it is the earlier shift of /u:/ > /y:/ which allows /u:/ to be a potentially better exemplar of /o:/ than /o:/ itself” (2004: 288). This seems to predict that no chains with the same steps but opposite chronological sequence would occur. Chains involving /u:/-fronting and /o:/-raising are actually quite common and have occurred in several series of

11 Crucially, note that in this model we are always discussing perceptual distance, not acoustic distance. I argue for the importance of this distinction in an earlier paper (Vaux and Samuels, 2004).
shifts, such as the Greek, Portuguese, and Scandinavian cases in Chapters 3 & 4. Though the literature typically calls these rephonologizations “push chains,” they clearly constitute drag chains if they occurred in the order for which Blevins has an explanation. Only establishing chronologies for these shifts through independent means can tell us whether they confirm Blevins’ implicit prediction or not. If the prediction is correct and \( /u:/ \) -fronting always precedes \( /o:/ \) -raising where it occurs, as long as there is some other way to spell out why \( /u:/ \) fronts all the way to \( /ü:/ \), Labov’s Pattern 3 shifts can be explained.

2.8 Synchronic chain shifts and modern functionalism

The chains discussed in this chapter thus far have happened over the course of generations, but linguists have also documented synchronic chain shifts. While analyses of these shifts do not bear directly on the issue at hand, there are important bits of information that can be gleaned from the literature on synchronic chain shifting.

In terms of rule ordering, synchronic and diachronic chain shifts both evidence counterfeeding\(^{12}\) opacity (Mortensen, 2004); the only difference is whether this is manifested in the synchronic or diachronic phonology of the language. Because counterfeeding opacity is a sensitive issue in phonological theory today, synchronic chain shifts have received a fair amount of attention in recent literature. At the heart of the issue is the contention that constraint-based Optimality Theory\(^{13}\) cannot handle opacity as well as a derivational theory of phonology can.

\(^{12}\) A *feeding* rule order occurs when the output of one rule serves as the input to a subsequent rule. The opposite of this is a *counterfeeding* order, in which the would-be feeder applies too late, i.e., after the rule that makes reference to its output.

\(^{13}\) Optimality Theory (Prince and Smolensky, 1993) was proposed as an alternative to traditional rule-based, derivational models of phonology. It replaces language-specific ordered rules with a set of violable and inherently conflicting universal constraints, which are ranked in a language-specific manner during acquisition and used in the production process to evaluate potential output candidates for a given input from the lexicon. OT constraints fall into two major types: markedness constraints, which penalize the
In order to pin down the exact nature of the opacity problem, some Optimality Theorists distinguish four types of synchronic chain shift mappings, which Lubowicz (2003) explains as follows:

(2.8.1) Typology of synchronic chain shifts (Lubowicz, 2003: 132)

(a) *push shifts*, with high-ranked markedness constraint against the initial stage in the shift but no high-ranked markedness against the intermediate stage,
(b) *pull shifts*, with high-ranked markedness constraint against the intermediate stage but no high-ranked markedness against the initial stage,
(c) *circular shifts*, with no termination point, and
(d) *regular shifts*, with high-ranked markedness against each stage in the shift.

Since the category of “regular shifts” as distinct from other push and pull chains is defined on a theory-internal basis, we will not concern ourselves with regular shifts any further. While circular shifts hold interest for phonologists because they are particularly difficult to motivate and cannot be captured very well by the markedness constraints typically employed in OT to account for chain shifts, they too fall outside the purview of the present discussion, which will be limited to push chains and issues directly pertaining to them.

Most OT approaches to synchronic chain shifting run into difficulty with push chains, just as many approaches to diachronic chain shifts do. This is perhaps unsurprising given the functionalist nature of OT constraints. Looking at a Natural Phonology account of diachronic chain shifting from the 1970s, one realizes that the notion of distinguishing push chains and drag chains by the “markedness” of their steps is occurrence of certain phonological elements, and faithfulness constraints, which penalize output forms that do not preserve the phonological identity of the input form in specific respects.
not so new after all. One linguist working in the then-new Natural Phonology framework wrote:

“Given chain shifts of the form A → B, B → C, C → D, etc., it is possible to determine whether one is confronted with a ‘push-chain’ or a ‘drag-chain’ by means of a fairly simple rule of thumb: it is a push-chain, if A → B is independently well motivated [i.e., relatively unmarked–BDS] while C → D is not understandable in abstraction from B → C; conversely it has to do with a drag-chain, if C → D has an intrinsic explanation, while A → B lacks any phonetic or structural motivation without taking B → C into consideration.” (Chen, 1974: 69)

I mention this intuition because it perfectly captures the inception problem as it pertains to push chains. Though Pattern 3 shifts happen in a C → D, B → C, A → B order, in this type of shift the C → D step (i.e., the initial [u]-fronting stage) is difficult to explain without positing some earlier encroachment from below (the B → C step). If we could only find a reason for the first step, we would solve the inception problem. If we were to co-opt Blevins’ theory in addition, we could explain Pattern 3 pattern shifts without a special mechanism of push shifting, and without encountering any of the criticisms that Martinet faced!

### 2.9 Interim conclusions

In this chapter we took a brief look at the Structuralist origins of chain shifting, beginning with Martinet’s formal introduction of the push chain/pull chain dichotomy in the 1950’s. We then saw how the debate over push chains is but one facet of a larger debate between functionalists and formalists in phonology, a nearly century-old battle about the nature of phonological systems and changes affecting those systems. After this discussion, we turned to two unsolved problems in historical phonology, both of which directly involve push chains: is sound change gradual, and does functional load play a role in sound change? We will return to both of these issues. Finally, we saw the two exemplar-based models that have been proposed recently in lieu of the feature-based, instantaneous model that, as we noted, cannot admit the phonemic encroachment
associated with push chains. These two exemplar-based theories differ in several ways. Labov’s (1994) theory explicitly allows push chains, though it cannot account for them in practice, and Labov believes there is empirical evidence that they exist. The particular case Labov has in mind, the Northern Cities Shift, will be the subject of Chapter 5. Blevins’ (2004) theory cannot admit push chains either, but it faces fewer internal problems with its pull chain mechanism. Finally, we briefly touched upon the phenomenon of synchronic chain shifting, which seems to lend credence to the feature manipulation theory of sound change. In the following three chapters, we will turn to the empirical evidence for push chains, so as to better understand what a theory of chain shifting should be able to handle, and to investigate whether any of the open questions about the nature of sound change can be put to rest.
3. **Chapter 3 – Greek & Portuguese**

In the first writings about chain shifts, André Martinet used the São Miguel dialect of Azorean Portuguese as his example of a push chain. Subsequent authors typically refer instead to a very similar shift in Attic-Ionic Greek, perhaps because the Greek case is better documented and additionally, Greek is more familiar to classically-trained linguists. In this chapter I will address Greek first, since it is possible to establish a solid chronology for the changes that took place in Attic-Ionic; for Portuguese we must place our trust in the ability of earlier linguists to write accurate descriptions of their observations and to draw the appropriate conclusions from those descriptions. The Greek and Portuguese starting points and subsequent changes are so similar, however, that it seems quite reasonable to explain them in the same fashion.

### 3.1 Prehistory: the Attic-Ionic vowel system

The Proto-Greek vowel system had five short vowels—two high, two mid, and one low, each with a long counterpart. This system essentially continued the Proto-Indo-European system; Greek short vowels represent PIE short vowels, while Greek long vowels represent PIE long vowels or short vowel + laryngeal sequences (Sihler, 1995: 35). I schematize the Proto-Greek vowel situation, which is also attested in Mycenaean inscriptions from the second millennium BCE, in (3.1.1) below.

(3.1.1) The Proto-Greek vowel system ca. 1800BC (Bartoněk, 1966)

\[
\begin{array}{ccc}
  i & u & i: \quad u: \\
  e & o & e: \quad o: \\
  a & a: \\
\end{array}
\]

We will be focusing on the Attic-Ionic dialect group. Proto-Attic-Ionic was probably spoken near Attica in the early first millennium BCE. Attic can be seen as
distinct from at least some varieties of Ionic by the early eighth century BCE, simply by virtue of Ionian colonization of the west coast of Asia Minor during this time period, which isolated what would become the East Ionic dialects. Specialists identify three dialect groups within Ionic, all of which were distinct by 500 BCE: West Ionic spoken in Euboea and north of Attica, Central Ionic spoken in the Cyclades, and East Ionic spoken in Asia Minor, south of Smyrna (Buck, 1965).

Unlike most of the other dialect groups, Attic-Ionic did not merge the results of compensatory lengthening and contractions with the inherited long mid vowels. They instead created another set of long mid vowels, higher than and distinct from the original two. These are known as the “spurious diphthongs,” orthographically represented by <ει> and <ου> in the standardized Attic system. The inherited [ej] and [ow] diphthongs merged with these vowels by the fourth century BCE in Attic-Ionic. The result of the new mid vowels was a system like the one in (3.1.2).

(3.1.2) The (early) Proto-Attic-Ionic vowel system, ca. 1000 BCE

\[ i \quad u \quad i: \quad u:\]
\[ e \quad o \quad \varepsilon: \quad \varphi:\]
\[ a \quad \varepsilon: \quad \varphi:\]
\[ a:\]

The Northwest Greek dialects, Megarian, Corinthian, and East Argolic also followed the Attic-Ionic pattern, though they differed slightly in the details of exactly what went into the “secondary” mid vowel category. Pamphylian, East Aegean Doric, 

14 The Boeotian adoption of the Ionic alphabet (probably via Attic) provides a terminus ante quem: the first half of the fourth century BCE. At that time, the Boeotians adopted <ου> for writing monophthongal [u] (Bartoněk, 1966: 80). The Boeotian borrowing will be important for dating the [u] > [ʊ] and [ʊ:] > [u:] shifts, and we will return to this fact shortly.
West Argolic, and Cretan had four height distinctions as well, but they derived this system in other fashions (Bartoněk, 1966. In particular see p. 61 for a chart summarizing the dialectal differences). Though symmetric, from a Structuralist perspective the four distinctions of height among the long vowels was an inherent source of instability for the group of dialects with this type of vowel system (Martinet, 1952, 1955; Ruipérez, 1956; Bartoněk, 1966; Szemerényi, 1987 [1968]). Over the first millennium BCE, the system of Attic-Ionic underwent numerous changes, some common to the whole Attic-Ionic group and others unique to a subset of the dialects.

Even prior to the first Attic inscriptions, Proto-Greek /a:/ raised and fronted, ultimately merging with /ɛ:/ in Attic and Ionic. Many linguists explain this as the first of many changes which began to alleviate pressure in the crowded back series of the long vowel system (Ruipérez, 1956; Bartoněk, 1963, 1966; Szemerényi, 1987 [1968]). Orthographic evidence supports the view that [a:] raised to [ɛ:] through the intermediate value [æ:]. Cycladic Ionic inscriptions directly attest a three-way distinction between [a:], [æ:], and [ɛ:]. Most notably, the well-known Nikandré inscription from the island of Naxos (late 6th century BCE) has etymological /*e:/ represented by <E>, etymological /*a:/ represented by <H>, and new /a:/ from compensatory lengthening represented by <A> (inscription and analysis available in Buck, 1965: 190; cf. also Gates, 1976: 190). From this one can conclude that [æ:] did not merge with [ɛ:] in Cycladic until sometime after the sixth century BCE. This interpretation is quite uncontroversial and appears in virtually the same form in the standard reference works on Greek, such as Allen (1974), Buck (1965), Sihler (1995), and Smyth (1956).

Of course, the Cycladic situation does not reflect on the date of the merger in the other Attic-Ionic dialects. Any history of Attic-Ionic must also explain why [a:] fronted in the first place. Let us set this issue aside for the time being, because in many respects [a:]-fronting parallels [u:]-fronting, which we will see in just a moment. The reader should note that the [a:] to [æ:] shift is yet another case of the inception problem rearing
its ugly head. It seems easy enough to explain the collapse of [æː] and [ɛː], but in order to get to the stage at which merger becomes possible, the [æː] must first develop for independent, as yet unidentified reasons.

Later, the loss of /n/ in –Vns- clusters resulted in lengthening of the vowels in these environments (Smyth and Messing, 1956: 14). This change is commonly known as the second compensatory lengthening.\textsuperscript{15} The result of <α> undergoing the second compensatory lengthening yielded a new [aː] phoneme, not [æː], and can thus be conclusively dated after the [aː] > [æː] shift. The second compensatory lengthening yielded uniform results in all Attic-Ionic dialects, so it must have occurred before the dialects split; this implies that the [aː] > [æː] shift happened during the Proto-Attic-Ionic. At this stage, according to conventional wisdom, it was the long front vowels that suffered from overcrowding.\textsuperscript{16}

After Attic had split off from Ionic, the remaining /æː/ merged with /ɛː/ in both Attic and Ionic independently, returning the system to its former, four-height-distinction, symmetric state. Evidence for dating the merger after the Attic-Ionic split comes from the

\textsuperscript{15} The first compensatory lengthening, also a Proto-Attic-Ionic change, involved loss of [h] (derived from Proto-Greek *s and *j) in certain sequences. The [aː] created by the first compensatory lengthening did feed into the [aː] > [æː] change, so it must have occurred prior to any of the other changes we are discussing.

\textsuperscript{16} An alternative interpretation, suggested by Jeremy Rau, holds that the introduction of new /aː/ from compensatory lengthening actually \textit{preceded and caused} the fronting of old /aː/. Since the new /aː/ was the result of denasalization, there was most likely an intermediate step in which the nasal consonant was lost but nasality remained on the vowel. This is supported by the eventual result of compensatory lengthening—nasalized vowels tend to be longer than corresponding oral vowels (Whalen and Beddor, 1989). This nasalization could have originally supported the old /aː/ ~ new /aː/ ([ãː]) distinction, but the distinction could have later been transferred to the front/back axis prior to or concurrent with loss of nasality. Thus, polarization of the two /aː/ phonemes would have resulted, with original [aː] becoming [æː] and previous [ãː] remaining further back. Note that such polarization can easily be explained as motion by both vowels towards perceptual equidistance, as would be predicted by the same mechanism in Blevins’ (2004) framework that accounts for (pull) chain shifting.
Cycladic inscriptions that show a three-way distinction, which precludes the possibility of a Proto-Attic-Ionic merger. Other evidence comes from the “Attic reversion,” which resulted in differing treatments of /æ:/ after {e, i, r} in Attic and Ionic. This complicated data does not bear directly on the issue at hand, so I will refrain from presenting it here, but I encourage the interested reader to consult this author’s recent paper (Samuels, 2005) for a detailed analysis of the situation.

This concludes the discussion of events prior to the chain shift that is of primary interest in this chapter. In subsection 3.2.1, we will discuss a more relevant aspect of the chronology presented in this section: when this merger of [æ:] and [ɛ:] occurred relative to the shifting of the back vowels. With this sequence of events securely dated, the case study can begin in earnest: that is, it will be possible to trace the developments of the [originally] back long vowels. These developments include the fronting of [u:] to [ü:] and the raising of [ǭ:] to [u:], a sequence which numerous introductions to historical linguistics cite as the flagship example of a push chain.

3.2 Evidence & chronology

3.2.1 The [æ:] ~ [ɛ:] merger

The Nikandre inscription, which maintained the distinctions between [a], [æ], and [ɛ:], provides incontrovertible evidence that [æ:] remained separate from [ɛ:] until the sixth century, at least in Naxos. For Amorgos and Keos, the same argument holds; there the three vowels remained distinct even into the fifth century BCE (Buck, 1965). This fact about the island dialects says nothing about the situation elsewhere in Attic-Ionic territory, however.

Bartoněk (1963; 1966), Allen (1974) and Ruipérez (1956) insist that the [æ:] ~ [ɛ:] merger transpired by 700 BCE, except in the dialects just mentioned, for which they assume that the merger had not yet occurred at the time of /u:/-fronting. Szemerényi (1987 [1968]: 1346), in direct opposition to the other scholars, retorts that “this is quite
impossible, and merely postulated on account of the ‘system’.’” He further notes that long
α-stem and η-stem nouns remained distinct until the end of the fifth century BCE, and that
choral lyric and choruses of tragedies perfectly repair instances of [æː] by lowering them
to [aː], without any confusion regarding what belonged to the [æː] class and what had
original long /e/. Therefore, Szemerényi follows the conclusion reached earlier by R.W.
Tucker (1962) and Edgar Sturtevant (1940), that the merger took place in Central and
West Ionic as well as Attic at various times throughout the fifth century BCE, while only
in East Ionic did it happen earlier, possibly as early as 700.

Early inscriptions from Attica and Ionian territory can easily settle this debate. In
even the most archaic texts from everywhere but the Cyclades, we see variants of epsilon
(<E>) representing etymological /aː/. Specifically, instances of <E> corresponding to
what we know to be older /aː/ appear in personal names in Attic graffiti from the late
seventh century BCE, and in Euboean inscriptions from c.550 BCE (Jeffery, 1961). If there
were still a distinct, phonemic [æː] in the non-Cycladic dialects, one would not expect
them to tolerate such orthographic conflation of [æː] and [ɛː]. Moreover, in the choruses
that Szemerényi mentions, incorrect “repair” of [æː] > [aː] sometimes occurred, which
suggests that the writers no longer maintained this class themselves (Björck, 1950). All
this evidence indicates that the merger of Proto-Greek /aː/ (i.e., Attic-Ionic [æː]) and /ɛː/
occurred prehistorically in all these dialects.

3.2.2  υ > [ʊ]

In 403 BCE, the Athenians officially began using the Ionic alphabet, and other
groups soon followed suit, abandoning their dialectal variants of the script in favor of the
standard version. The Ionic alphabet was primarily characterized by the use of <η>,
indicating a long [ɛː] sound instead of its former phonetic value of [h] (though Ionic
shared this feature with some other scripts), and <ω>, innovated for symmetry, indicating
a long [ɔː] sound (Buck, 1965: 19). Among the groups to begin using the Ionic alphabet
in the early fourth century BCE were the Boeotians. When the Boeotians adopted the Ionic
system, probably through an Attic intermediary, they used the spelling <ου> for their [u] sound. This indicates that the phonetic value of <ου> in Attic-Ionic by the early fourth century was more similar to [u] than the value of <υ> was (Bartoněk, 1966: 111). Attic-Ionic /u/ must have fronted to [ü] by this time, though the question of whether fronting occurred in Euboean (either it did not, or a reversion took place) remains open (see Bartoněk, 1966 for one point of view; del Barrio Vega, 1990 and; Méndez Dosuna, 1993 for the other). Whatever the case, Euboean did not participate in any of the other chain shift steps.

Another piece of orthographic evidence that originally led linguists to recognize /u:/-fronting in Attic was once seen as compelling, but can now be rejected. The letter <Ϙ>, known as koppa or qoppa, was a letter with the phonetic value [k] that appeared in early Greek inscriptions instead of kappa when preceding the sounds /a/, /o/, or /u/. Linguists noticed that the sequence <ϙυ> is poorly attested in Attic-Ionic, while <ϙο> spellings were once relatively common. This led some to believe that /u/ had fronted in these dialects by the time of the earliest inscriptions. Furthermore, there is the odd case of an Athenian amphora from circa 570 BCE with Κϙυλωιος written on it, as if the scribe could not decide whether kappa or qoppa would be more appropriate before upsilon (Threatte, 1980: 22-23). This, too, has been taken as evidence for fronting of /u/ around that date. The value of this data becomes less clear when one realizes that the first attestation of <κυ> is from about 550 BCE, by which time <κο> was gaining ground against <ϙο> and there was a “prevailing tendency to liquidate completely the sign φ” in favor of simply using kappa across the board (Bartoněk, 1963: 32). Thus, fronting might not have taken place this early. On the other hand, Bartoněk (1963: 33) notes that even a fronted [ü] was “no doubt sufficiently rounded to admit phonetically of its eventual

17 Bartoněk (1963; 1966) in fact claims that this sequence is never attested in Attic, but Threatte (1980) provides evidence to the contrary, including two cases of <ϙυ> in graffiti dated to 600-575 BCE.
graphic combination with the foregoing ϕ.” Being no ancient Athenian myself, I am far less confident on this point, but it certainly merits consideration. Méndez Dosuna (1993) goes even further, saying the distribution of kappa and koppa must be considered completely arbitrary, a mere vestige of the Phoenician system from which it was borrowed. This seems overly pessimistic, though at any rate, it is most prudent to conclude that the data pertaining to koppa is not at all as clear as it was once made out to be; it would be best to look elsewhere for corroborating evidence.

With the help of other orthographic evidence, we can ascertain the date of /u:/-fronting with greater precision. When the conventional spelling of the diphthongs was established around 700 BCE, /u/ must still have been a back vowel, or <EY> and <OY> would have failed to make sense as spellings for [ew] and [ow], for which all indications point to a back offglide (Bartoněk, 1966: 114). By the sixth century BCE, some Ionic texts show <EO> and <AO> rather than the conventional spellings for the diphthongs (Bartoněk, 1966: 113). This evidence is somewhat ambiguous, but it may suggest that /u:/-fronting had taken place by that time, leading to confusion about the identity of the diphthongs’ offglide. One also finds confusions of /u:/ and /i:/ by the sixth century, which can only be explained by /u:/ being realized as a front vowel at that date (Teodorsson, 1974: 289). Allen (1974) makes clear, however, that this does not necessarily mean that the pronunciations of <υ> and <ι> had already merged, as they eventually did, producing modern Greek [i]. What we can deduce from the combination of all this evidence is that the fronting of [u:] to [ü:] occurred between 700 and 600 BCE.

3.2.3 υυ > [u:]

Any analysis of the Greek situation crucially depends on the timing of the /u:/-fronting step relative to the /ϕ:/-raising step. If [ϕ:] raised (partway) to [u:] prior to or concurrently with /u:/-fronting, that would be evidence for a push chain in the truest sense: an encroachment (the raising of /ϕ:/) causing displacement of another phoneme (original /u:/). This would have important theoretical ramifications: there would be no
choice but to allow teleology as an explanation and also accept the gradualness of sound change. We would also be left without a theory, given that no proposed mechanism can capture shifts that involve genuine encroachment. Thankfully, as it turns out, there is good reason to believe that /u:/-fronting preceded /o:/-raising by several generations, and therefore this step was certainly a pull, not a push.

From the fourth century BCE onward, no confusions of <ω> and <ο> exist in Attic texts (Teodorsson, 1974: 213). Examples of <ο> substituted for <ου> also become rare after about 360 BCE in Ionic (Lasso de la Vega, 1956: 280) and 325 BCE in Attic (Threatte, 1980: 258), setting aside cases in which <ο> was most likely an abbreviation for the genitive ending <ου>, which is a well-documented practice. This suggests that the vowels were increasingly differentiated by the raising of the formerly close mid long vowel around this time. There is also a parallel change in the front vowels that would fit nicely in this chronology: a change which most likely originated in Boeotian, /ɛ:/ raising to and merging with [i:], occurred sometime in the third or fourth century BCE in Attic and also happened during the latter part of that time frame in Ionic (Bubeník, 1983). The raising of the two close mid vowels would then have operated almost concurrently, a situation leading to mutual reinforcement.

It appears that the Attic-Ionic raising of /o:/, then, transpired during the fourth century BCE. However unlikely it seems that “the opportunity would long have been resisted of increasing the acoustic distance” between the two mid back long vowels after /u:/ fronted (Allen, 1974: 77-78), one cannot escape the conclusion that the two changes were separated by centuries. For this reason, I agree with Szemerényi (1987 [1968]: 1349) that “it seems safest… not to force the notion of a push-chain on these changes. We should rather view the (later) change ρ > ū as being the exploitation of an existing possibility, namely of the empty square.”

I have already mentioned that, if this series of changes can be viewed as a push chain at all, the “push” would have to be overcrowding, not encroachment. Virtually all
analyses accept the relative chronology I have given with /u:/-fronting preceding /o:/-raising, even if they do not agree with the absolute dates I propose. In over a year of research, I found only one serious consideration of a chronology in which raising pre-dated fronting, and even that source (Bubeník, 1983) did not fully espouse it.

With the Greek chronology established, we now turn to a situation that is exactly parallel, but for which the evidence is of an entirely different nature. Once we have investigated this shift in Portuguese and pursued the matter of another similar shift in Old Scandinavian in Chapter 4, we will be in a position to discuss the very weighty question of whether cases like these should rightly be considered push chains.

3.3 A parallel situation: São Miguel Portuguese

The dialect of Portuguese spoken on the most populated of the Azores, São Miguel, is distinguished from the other Azorean dialects, Madeiran, and standard Lusitanian by several phonological features. In 1939, Francis Millet Rogers, then a graduate student in comparative philology at Harvard, spent a month collecting linguistic data on São Miguel. This work, along with his other travels in Portuguese-speaking territory, provided the basis for his doctoral dissertation (Rogers, 1940) and a subsequent article on insular Portuguese pronunciation (Rogers, 1948). André Martinet cited Rogers’ 1948 article, in particular the discussion of the unusual São Miguel vowel system, in his discussion of chain shifting. As we have already noted, Martinet (1952) provided the São Miguel case as his lone example of a push shift; Haudricourt & Juillard (1949) also mentioned the dialect in their list of systems—French, Greek, and Swedish being the others—that exhibit /u:/-fronting. To add to Rogers’ data, we now have a small sociolinguistic study recently undertaken by David J. Silva. For the most part, Rogers and Silva are in agreement about são-micaelense.

In his dissertation, Rogers (1940: 432-433) identifies fifteen “general phonetic characteristics” of the São Miguel dialect. With respect to back vowels, he notes that where standard Portuguese exhibits stressed [u], são-micaelense instead shows [ü], even
for well-educated speakers and in careful speech. This shift occurred not just in São Miguel, Rogers notes, but also on the south coast of Madeira, on Corvo, and in some continental dialects. Standard Portuguese [o] also quite reliably fronts to /œ/ (somewhere between [œ] and [ø]) in são-micaelense, but only when spelled <ou>. The standard Portuguese [o], spelled <ô>, is fairly often realized as [u] on São Miguel. This led Rogers to postulate that there were two variants of /o/ in Portuguese at the time São Miguel was settled, and that these two vowels merged in the standard dialect but remained distinct in são-micaelense. The standard [ɔ] (<ô>) is often pronounced [o], with what Rogers notates as [o] in the standard dialect (i.e., IPA [a]) ranging from [a] about half the time to higher and backer variants [a] or even [ɔ]. Even more sporadically, the front vowel [e] may be realized as [ɛ], and [ɛ] can lower as far as [æ]. I schematize both the standard and São Miguel vowel systems in (3.3.1) below.

(3.3.1) Standard Portuguese and São Miguel Portuguese stressed vowel systems

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>u</td>
<td>i ü</td>
<td>(u)</td>
</tr>
<tr>
<td>e</td>
<td>o</td>
<td>e ø</td>
<td>o</td>
</tr>
<tr>
<td>ɛ</td>
<td>ɔ</td>
<td>ɛ</td>
<td>ɔ</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>(æ)</td>
<td>a</td>
</tr>
</tbody>
</table>

Standard Portuguese

Parentheses indicate variants not present in all speech communities. São Miguel

[ɛ] stems from the merger of original [e] and optionally monophthongized [ej];
[øj] and [ow] monophthongized to [o] are also accepted.

---

18 Rogers (1940: 345) says of Madeiran /ü/: “This vowel is pronounced more or less like the (y) of French pur, although slightly less ‘forward’,” and that the /ü/ on São Miguel is, if anything, more like the French vowel. He also notes that other scholars say these vowels resemble the Swedish & Norwegian /u/ (1940: 349).
When Silva visited São Miguel in 1985, he found much the same situation as Rogers had described. He conducted and recorded interviews with twelve residents of Nordeste, a village on the northeast coast of the island, and ran statistical analyses of their pronunciation, which he reports in his 2005 *Luso-Brazilian Review* article. Totaling up the instances of /u/ in the interviews, Silva discovered that 100% of them were fronted to [y], “even in the most careful speech” (2005: 11). Orthographic <ou> became [ø] 96% of the time, standard ç was pronounced as [a] 88% of the time, and /a/ was realized as [a] or [ç] in 42% of tokens. Both the backing of [a] and, to an even greater extent, the lowered realizations of the front vowels were so infrequent as to be “dispreferred” according to Silva. These numbers seem to reflect Rogers’ impressionistic ideas about the extent to which each of the shifts had progressed. One seeming anomalous piece of data in Silva’s analysis is that orthographic <ô> was pronounced [u] only 38% of the time. For some speakers, Silva reports, the two mid back vowels have merged, rather than the higher of the two raising to [u]. This has the net effect of reducing the vowel inventory in this dialect by one phoneme.

Unfortunately, this is the extent of the data on the São Miguel back vowel shift. In terms of drawing conclusions about the ordering of the links in the chain, we can only rely on the well-accepted principle dating back to Bloomfield (1933) that linguistic change proceeds like an S-curve (see Labov, 1994 for a bibliographic history). A change in pronunciation (or syntax, or morphology, or any other linguistic variable) picks up steam gradually at first, then gains ground quickly, and finally tapers off as it nears completion, as shown in the graph in (3.3.2).
From this, we can conclude that a hypothetical alternate pronunciation occurring in 30% of tokens stems from a newer change than one occurring in 70% of tokens. Using this principle as a guide, we establish [u] > [ü] as the oldest change, <ou> > [ø] as the next,\(^{19}\) [ɔ] > [o] somewhat later, and [a] > [a] ~ [ɔ] as fairly recent, with the lowering of [ɛ] and [ɛ] still in the early stages of progression. For <ô>, I propose that, in the segment of society for whom [o] > [u] occurred fairly early, there was no merger of /o/ vowels, whereas in the segment for which [o] remained for a longer time, the merger took place. The chronology of shifts the non-merging dialect, it is interesting to note, looks exactly like an extended version of the Attic-Ionic chain shift.

### 3.4 Analysis

Since both the Attic-Ionic and São Miguel shifts appear to have started with /u/-fronting, it is critical to explain that step in order to claim understanding of the changes that took place in these dialects. The difficulty in explaining /u/-fronting lies in the fact that, in both cases, there were other dialects of the same language that had the same four-

---

\(^{19}\) This would agree with Labov’s observation that “the fronting of /o/ has no connection with either chain shifting or the principles of chain shifting. It is never linked directly with the raising of a back vowel that would otherwise merge with the mid vowel, but instead represents a generalization of the fronting of the high back vowel.” (1994: 208).
height-distinction vowel system, yet did not undergo the same changes. Szemerényi comments on this problem:

“Although the development we are discussing has been stated almost in terms of a natural law in the case of vowel-systems with four degrees of aperture in the back row, there have also been many dissentient voices. Thus, e.g., in Gallo-Romance Lat. ū developed into ü, while in Tuscan we find no such change. Similarly, there were a number of Greek vowel-systems with four degrees of aperture—but only in Ionic-Attic did ū develop into ü.” (1987 [1968]: 1348)

For this reason, some linguists have called upon extra-linguistic factors to explain /u:/-fronting in both Greek and Portuguese. Of Attic-Ionic, Bartoněk says that /u:/-fronting “required the coincidence of two factors, first of the substratum impulse from Asia Minor, and secondly a favourable systemic condition implying the overloading of the back long-vowel axis,” and that “this coincidence appears to have asserted itself at the time when the substratum influence of Asia Minor was still strong enough only in the Attic-Ionic dialects” (1963: 38-39). In the same vein, early observers of insular Portuguese attributed the fronting of /u/ to “Keltic” substrate influence—as scholars trying to explain French /u/-fronting have also done (cf. Haudricourt and Juilland, 1949; Malmberg, 1964; Ahn, 2004)—because of circumstantial evidence pointing to a small Breton community on São Miguel. Rogers (1940) dismisses this idea, saying such a small group was extremely unlikely to prove so influential. Because we cannot see into the past, we cannot definitively rule out even the most unlikely substrate (or neighboring language) influence, but it is unwise to appeal an implausible historical scenario before exploring all other avenues; I agree with Malmberg (1964: 19) that “the substratum explanation has been seized on far too frequently as a convenient solution and often on very slender grounds.” One should at least try to find other factors. In this case, thankfully, these are very possible to find.

In his recent book, Andrea Calabrese (2005) reiterates a point most explicitly defined by Optimality Theory practitioners (cf. §2.8) but which is an implicit
undercurrent in the earlier literature describing these particular shifts: the fronting of [u] to [ü] is an instance of “emergence of the marked.” In other words, [u] is a much more frequently occurring (“unmarked”) sound than [ü] in the world’s languages, and the change of the relatively unmarked sound [u] to the more marked sound [ü] is an unexpected development. All else being equal, linguistic change typically eliminates markedness rather than creating it. There must be some overriding factor causing these Pattern 3 chain shifts, something powerful enough to overcome the disparity in markedness between [u] and [ü].

The fact that we cannot isolate any necessary and sufficient causes for /u/-fronting means that, whatever our theory of why this change occurs, it will not have perfect predictive value. Like researchers studying a disease with both genetic and environmental components, the best we can do is identify “risk factors” that predispose a language towards developing in this particular fashion. At this point, having examined two very similar shifts, we can begin to identify some of these risk factors.

Clearly, Martinet and numerous subsequent theorists have noted, having four distinctions of height among the back vowels predisposes a language towards /u/-fronting. Labov states that “the full chain shift we call Pattern 3… takes place only when there are four degrees of height in back. The two examples of vowel shifts in Martinet (1955) and the further cases explored by Haudricourt & Juillard [French, São Miguel, Greek, and Swedish] are confined to this pattern” (Labov et al., 1972: 105). Twenty-five years after that original statement, Labov himself came up with a counterexample: the American West, which exhibits both a three-height distinction in the back of the mouth owing to the merger of the vowels in caught and cot. Despite this, like virtually all other American English dialects, this dialect fronts /u/ (Labov et al., 1997). Also, the seven-

---

20 This is but one simple formulation out of the many possible ways to define markedness; there is no one definition agreed upon by all phonologists. By any set of criteria, though, [ü] is marked relative to [u] (Calabrese, 2005: 47).
vowel system with three height distinctions among the front vowels (i, e, ε) and four among the back vowels (u, o, ɔ, a) happens to be one of the two most common seven-vowel systems across languages (Crothers, 1978). This would seem to argue against any proposal that argues for the inevitability of /u/-fronting simply based on the fact of a four-height distinction of back vowels, though of course frequency is not a guarantee of stability.

Another factor that has been discussed is the lack of high central vowels in /u/-fronting languages. For Attic, Teodorsson (1974: 291) posited a “latent attraction” of /u:/ to /i:/ owing to the lack of any central vowel between them. Since the Attic situation was not resolved by /u:/ merely centralizing, and because there is no reason to believe that languages prefer having a high central vowel to having a high back vowel, one can conclude that any “attraction” would be concerned with more than just filling the gap between the high vowels. At this juncture one could invoke the Pierrehumbert and Blevins approach, in which vowels move towards empty spaces to minimize perceptual confusion. As noted earlier, however, this principle alone certainly cannot explain why /u/ would move all the way to the front. It also fails to explain why /u/-fronting would occur (though of course on a much smaller scale) in languages such as English, which have competing central vowels. One could make the argument that the perceptual distance between the high central and high back vowel before fronting is still greater in these cases than the distance between the high back vowel and any other neighbor; this ought to be experimentally verifiable.

Calabrese (2000; 2005) fights fire with fire, proposing that the force overcoming the markedness of [ü]/[y] is in fact another markedness constraint. He builds on the work of Lindau (1975) and Archangeli & Pulleyblank (1994) to establish an articulatory account of why [u] and [ɔ] are articulatorily complex in a way that [y] and [ø] are not. The vowels [u] and [ɔ] are both pronounced with the tongue root in an advanced position (in featural notation, this is represented as [+ATR]). As Lindau first noticed, advancing
the tongue root pushes the body of the tongue up and forward, which means that “in pronouncing [+ATR] back vowels, speakers need to suppress the natural tendency to front them” (Calabrese, 2005: 49). Thus, the changes of [u] > [y] and [o] > [ø] simplify articulation by preserving the [+ATR] feature at the cost of losing the [+back] feature. In Calabrese’s framework, this is manifested by a constraint disallowing [+ATR, +back] vowels with an associated repair deleting [+back] when [+ATR] is also present. Of course, one certainly need not subscribe to this particular theoretical idiom in order to make the best of the observation. It would suffice to say that the articulatory complexity of [u] and [o] make them susceptible to fronting.

Two other details about ATR are worth mentioning. Firstly, the low [+ATR] vowel /A/ in Somali is realized as [æ] in a process that also produces [y] from underlying /u/ and [ø] from underlying /o/ (Calabrese, 2005: 50). This is a synchronic process of phonetic realization, but it is certainly tempting to draw a parallel between the Somali case and the Greek diachronic processes of [a:] becoming [æ:] and [u:] fronting as well. One could then view /a/-fronting and /u/-fronting as (partially) stemming from the same cause. Secondly, Calabrese (2000; 2005) reports that, in the Altamura dialect of Italian, [+ATR] rounded vowels front only when they appear in stressed syllables. This is reminiscent of São Miguel Portuguese, in which the readjustment of vowels, including the fronting of /u/ and <ou>, primarily occurs under stress. This does not necessarily imply a connection between stress and ATR in particular, but it provides another instance of stress interacting with vocalic fronting.

In this chapter, we have seen two cases of chain shifts beginning with /u/-fronting and we have identified several factors that appear to create circumstances favorable to /u/-fronting: four height distinctions among the back vowels, a lack of high central vowel, and the articulatory complexity of the [+ATR, +back] combination. Though this goes a long way towards explaining how Pattern 3 shifts begin and partially answers the “general problem” (Coseriu, 1958) of determining the conditions under which such chain shifts begin, there are still many questions about chain shifting that we have yet to
answer. On our mission to find those answers, we will next turn to the Old Scandinavian vowel shift, which shares some features with the Greek and Portuguese cases but also differs from the shifts we have already seen in interesting ways.
4. Chapter 4 – Swedish & Norwegian

4.1 Prehistory: the Old Scandinavian vowel system

The next case study involves two closely related Germanic languages, which survive up to the present as Swedish and Norwegian. Just one thousand years ago, they were mutually intelligible dialects of a language that linguists call Old Scandinavian, Old Nordic, or Old Norse. Some of the early developments discussed in this chapter are common to Swedish, Norwegian, and their close relative, Danish, while others only occurred in certain areas of the linguistic territory in question. We will first begin by tracing the evolution of Old Scandinavian into its various daughter languages, so as to better understand the historical context of the chain shift we will be investigating.

The earliest texts from the Germanic family are runic inscriptions, the oldest of which date back to the first century CE (Fortson 2004: 310). These inscriptions already show features common to the modern Scandinavian languages but not to East Germanic (a now-extinct branch that included Gothic, Vandalic, and Burgundian) or West Germanic (represented by living descendants including English, German, and Dutch). For this reason, many linguists view the language of the earliest runic inscriptions as the direct ancestor of the Scandinavian/Nordic or North Germanic group (Haugen, 1976: 107). In the centuries during and after the writing of the runic texts, Scandinavian speakers migrated far and wide. The Vikings, famed for their raids and seafaring exploration, helped with this linguistic expansion throughout the period from 800-1050 (Haugen, 1976: 134). While all this transpired, the language of the Vikings, commonly known as Old Norse, gradually split into different languages (Fortson, 2004: 328). By 1100, one can identify several distinct dialects: Old Norwegian and Old Icelandic, belonging to the West Scandinavian group, and Old Swedish, Old Gutnish (spoken in

21 Old Norse is also sometimes used as a synonym for Old Icelandic, but here it should not be interpreted as such.
Gotland), and Old Danish, belonging to the East Scandinavian group (Haugen, 1976: 190).

The early Old Norse vowel system, attested in the first runic inscriptions, had five short vowels and five long vowels. I schematize this system in (4.1.1) below.

(4.1.1) The early runic vowel system ((adapted from Faarlund, 2004: 911)

\[
\begin{array}{cccc}
i & u & i: & u: \\
e & o & e: & o: \\
a & & & a:
\end{array}
\]

This represents only a slight change from the Proto-Germanic system: one Proto-Germanic long /e:/ (sometimes called ē₁) became /a:/ in North Germanic, while the other (ē₂) became North Germanic /e:/ (Faarlund, 2004: 911). Between the period represented in (4.1.1) and the ninth century, Old Norse underwent many phonological changes, one result of which was a much expanded vowel inventory. The newly incorporated vowels were the front rounded [y] and [ø], plus [æ] (or a similar low front vowel) and [ɔ]; each of these came in both long and short varieties (Haugen, 1976: 196; Fortson, 2004: 330). The later Old Norse system, then, came to look like the one illustrated in (4.1.2).

(4.1.2) The later Old Norse vowel system

\[
\begin{array}{cccc}
i & y & u & i: y: & u: \\
e & ø & o & e: ø: & o: \\
 & & & & \\
 & & & & \\
 & æ & a & æ: & a:
\end{array}
\]

\[\text{22} \text{ The [y] was the result of i-umlaut of [u], [ø] came from u-umlaut of [e], [æ] came from i-umlaut of [a], and [ɔ] came from u-umlaut of [a].}\]
This sight, a vowel system with four height distinctions in the back, should by now seem rather familiar. In approximately the ninth century, with the vowels arranged in this fashion, the stage was set for the changes collectively known as the Old and Middle Swedish or, more popularly, Central Scandinavian vowel shift (Riad, 2005: 1105).

4.2 Evidence & chronology

One important change in early Scandinavian phonology was the lengthening (also known as “the great quantity shift”) of short /a/ in open syllables. The early date of this change is evidenced by its uniform application and broad geographical distribution, covering “such otherwise different areas as Iceland, Midland Norway, Denmark, and Finland” (Haugen, 1970b: 133-134). On the other hand, orthographic evidence for the lengthening of [a] does not appear before 1250, even by the most liberal estimates (Wigforss, 1978). It is certainly not clear, then, that this was a very old Common Scandinavian development. The relationship of the lengthening to the Central Scandinavian shift proper is even less clear. As Riad (2005: 1106) notes, “there is no necessary connection” between the lengthening process and the chain shift proper, but they are often mentioned in conjunction with one another. Moreover, the timing of the lengthening affects the possible analyses of the chain shift; as we will soon see, the chronology has a couple different possible interpretations.

The Central Scandinavian shift, unlike the others discussed in this paper, involved a merger. Specifically, long /a:/ merged with /ɔ:/, resulting in a vowel that “may have

---

23 This shift affected all short vowels, not just /a/, and similar shifts occurred in most of the Germanic languages. The general tendency was for all stressed syllables to become heavy, through the lengthening of short vowels in open syllables and the shortening of long vowels in superheavy closed syllables (Riad, 2005).

24 This is a slight simplification of the facts; differing treatments of various consonant clusters show that lengthening must have actually proceeded in at least three phases, the earliest of which happened c. 1250, with subsequent phases completed by the early fourteenth century (cf. Wigforss, 1978).
been phonologically identical with $\phi$ [i.e., $\emptyset$ – BDS] or with $a$, or that... may have been intermediate between them, as far as its rounding was concerned, or... may have fluctuated between them in free or conditioned variation” (Benediktsson, 1970: 106).

Whatever its exact phonetic implementation, this merger certainly began prior to 1200, since Old Danish and Old Icelandic also share this development (Haugen, 1976; Wigforss, 1978). Orthographic evidence, namely the use of <aa>\(^{25}\) or <o> rather than <a> to represent the merged vowel, reveal that the merger had become widespread by 1200 in Norwegian and Icelandic, during the thirteenth century in Danish, and by 1350 in Swedish (Haugen, 1976; Riad, 2005).

As it turns out, the [a:] resulting from lengthening of short [a] did not participate in the merger with [œ:]. One can interpret this fact in a couple of different ways— Riad (2005), following Widmark (1998), claims that the lengthening of [a] in open syllables occurred before the merger of [a:] and [œ:] (cf. also Benediktsson, 1970), whereas Haugen (1976) and Wigforss (1978) argue for the reverse ordering. Neither of these competing scenarios can be excluded *a priori*, but I am inclined to trust the orthographic evidence, which seems to indicate that the raising occurred prior to the introduction of new long [a:] in Norwegian, Icelandic, and possibly Danish, but that it may have concomitantly with the lengthening in Swedish.

As uncomfortable as it may seem to reach the conclusion that the different languages may not have experienced the same ordering of steps, I do not feel this is problematic. We need not look to the lengthening to explain the raising or vice versa. Similar processes backing and rounding long /a/ took place in Middle Low German and English (Haugen, 1976: 257), and Old English also participated in a mild form of the quantity shift, “as part of a general trend in this direction of all west European languages”

\(^{25}\) An <aa> spelling could also just indicate vowel length, so it provides less conclusive evidence than an <o> spelling (Riad, 2005).
These changes, then, were not the result of any acute imbalance peculiar to Scandinavian, but should rather be seen as manifestations of much broader trends gradually diffusing over large swaths of Europe.

Furthermore, the net effect of the lengthening and the raising on the long vowel system (in any order) was zero. Both /a:/ and /ɔ:/ were present in the system prior to any changes, so no phonemes were created or lost in the long vowel system, nor was there any “encroachment” on the higher back vowels during the course of the lengthening and merger. Calling either of these early events “triggers” for the later chain shift, as some linguists have done (cf. Haugen, 1970a; Hock, 1986; Labov, 1994; Eliasson, 2005), therefore seems a bit of a stretch. Of course, it is still worth noting that after the dust had settled, the system still had four different height distinctions among the back vowels. We have already identified this as a risk factor for /u/-fronting, which is precisely what happened next in some parts of Scandinavia.

The geographical scope of /u/-fronting also extended beyond the core Central Swedish and East Norwegian dialects; other Swedish and Norwegian dialects also participated (Haugen, 1976). Some have also seen a connection between the Faroese, West Norwegian, and Gutnish diphthongization of /u:/ and the fronting of /u:/ in the other languages (Haugen, 1970a; Tomas Riad, p.c.). In its phonetic details, the Scandinavian /u/-fronting differs from the cases we discussed in Chapter 3. Since [y:] already existed in the phonemic inventory, the old [u:] did not shift all the way to the front. Instead, it became a high central rounded vowel, distinguished from [y:] by its central location and by the unique “outrounding” (i.e., rounding of spread lips) that [y:] developed (Eliasson, 2005). The date of this change is unfortunately not clear, so opinion has traditionally been divided as to whether it preceded or followed the raising of [o:].

Opinions differ as to what this connection might be. Haugen (1970a: 69) views /u/-fronting as “resistance to new diphthongization of the long vowels,” while Braunmüller (2003) believes that diphthongization was the primary method of change, even in the canonical Central Scandinavian shift.

---

26 Opinions differ as to what this connection might be. Haugen (1970a: 69) views /u/-fronting as “resistance to new diphthongization of the long vowels,” while Braunmüller (2003) believes that diphthongization was the primary method of change, even in the canonical Central Scandinavian shift.
The [o:] vowel—*not* the [ɔ:] vowel—raised to [ʊ:] sometime around 1400, according to Haugen (1976). Some scholars, including the early researchers Amund Larsen (1886; 1926) and Johan Storm (1881; 1908), believed that this change was precipitated by the earlier change of [a:] raising to [ɔ:]. I agree that the raising of /a:/ did take place prior to the raising of [o:], but I see no connection between these two changes given the fact that old [a:] merged with the already extant /ɔ:/, as I discussed earlier. The only possible case one could make would have to hinge on redistribution of functional load leading to a need to separate [ɔ:] and [o:], but given the highly subjective nature and lack of predictive value of functional load, such an account would be highly unsatisfactory. I believe it is better to divorce the lengthening of [a] and the raising of [a:] from the later two changes, which find ready explanations from other sources.

This leaves us with the question of how exactly to explain the movements of /o:/ and /u:/? Specifically, we must ask whether the raising of /o:/ pushed /u:/ forward, or whether the fronting of /u:/ dragged /o:/ higher. Eliasson (2005: 1120) gives several reasons why he believes the raising preceded the fronting, but none of them actually make the case. First, he cites the “internal logic of the constituent changes,” but if I am correct in keeping the early changes separate from the later ones, this argument is not tenable for reasons we have already discussed. Secondly, he appeals to the geographic distribution of the changes. We agree that geographical distributions are useful for establishing relative chronologies, and that in general a wider distribution means an older change. Eliasson does not, however, supply any evidence suggesting that the [] > [ʊ:] change is as widespread as the /u:/-fronting, nor have I seen any elsewhere. The third point Eliasson makes is that “certain philological evidence” suggests that the [ɔ:] > [ɔ:] change took place “comparatively early.” Again, we agree on this point, but I do not believe it has any bearing on the question of whether /u:/-fronting happened before or after /o:/-raising. Finally, Eliasson argues that “the unique structural exploitation of the
articulatory gesture of outrounding… can be readily understood under a push, but not so easily under a drag, interpretation.” I see no reason why this should be the case. The outrounding certainly does seem to be a response to the new coexistence of two non-back, high, rounded vowels, but one can easily explain the fronting of /u:/ that created the more central of the vowels without appealing to a push from below. The mere fact that there was clearly some strong pressure leading [u:] to front does not give any clues about what that pressure might have been, and it certainly does not necessitate upward pressure from the lower back vowel. In the absence of compelling evidence to the contrary, then, I interpret Scandinavian /u:/-fronting and /o:/-raising as yet another Pattern 3 shift of exactly the same type as the ones in Chapter 3.

4.3 Analysis

Though the Central Scandinavian shift has been recognized since the late nineteenth century (Storm, 1881; Larsen, 1886), questions about the exact nature and causality of the chain still remain unanswered. Einar Haugen, in a reply to a paper read by Hreinn Benediktsson, sums up the problem well:

“It would be folly to deny that it is systematic, in the way that has been observed since Amund B. Larsen’s time, and in a way that resembles that of the English vowel shift. But any attempt to connect systematicity with causality is bound sooner or later to face the problem: why do like causes not produce like results? A system which in its essentials was identical from Iceland to Finland and from North Cape to the Ejder River in fact went through as many different structural developments as its state permitted. The qualitative structure remained largely unaffected in such otherwise different areas as Iceland, Midland Norway, Denmark, and Finland. Yet all of these also took the initial step of lengthening the short a which is listed by Hreinn as the inciting moment of the change. All of them labialized and backed the long a which was the first step of the push-chain movement. Yet it was only a coherent area consisting of Eastern Norway and Sweden in which the vowel system as a whole shifted upward and forward in the odd way which still characterizes standard Norwegian and Swedish. I do not believe that we as linguists are in a position to answer this question, since the impulses are clearly extra-linguistic. No
structural or generative analysis can possibly give us the basic answers we are seeking.” (Haugen, 1970b: 133-134)

This is a perfect example of the “actuation problem” (Weinreich et al., 1968), which I mentioned in Chapter 1. In this case I believe the problem, as defined by Haugen, is actually illusory: the lengthening of [a] is a red herring. This change, and the merger of /a:/ with /ɔ:/, are entirely independent from the later chain shift.

On the other hand, an actuation problem of a different type still remains. After the changes involving the lower vowels, the systems were indeed the same—and had four height distinctions in the back—over the vast territory Haugen mentions. I still feel justified in asking why, of all these languages and dialects, only Swedish and Norwegian participated in /u:/-fronting and /o:/-raising. In Chapter 3, we encountered this same issue in Greek and Portuguese. The conclusion is inescapable: identical systems simply do not evolve in identical ways. Linguistic change has an element of randomness, even though it is possible to identify common types of phonetic/phonological, syntactic, morphological and semantic changes. The predictive value of any diachronic theory is correspondingly limited, but this does not mean we have to give up entirely. In fact, in Chapter 3 we identified commonalities between Greek and Portuguese that seemed to contribute to the Pattern 3 shifts in those languages, and Common Scandinavian also shared these characteristics. This is further evidence that we are correct to suggest that having four height distinctions in the back vowel system and lacking a high central vowel are risk factors for Pattern 3 shifting; the complexity of [+ATR, +back] vowels may also contribute to the fronting of /u/.

4.4 Interim conclusions: categorizing Pattern 3 chain shifts

In this chapter and the previous one, we have discussed three textbook examples of chain shifts involving /u/-fronting and subsequent raising of the close mid back vowel. I hope I have persuaded you of a fact that will be critical to our ultimate analysis of the push chain controversy: in none of our case studies did /o/ raise prior to the fronting of
/u/, but instead, /u/ fronted and dragged /o/ behind it. This particular type of shift is very common—so common that Labov (1994) christened it with a special name, Pattern 327. Examples of Pattern 3 shifts come from such varied languages as French (Haudricourt and Juillard, 1949), Armenian (Vaux, 1992), Somali (Antell et al., 1973), Swiss French, Yiddish, Lithuanian, Albanian, West Syriac, Akha (Lolo-Burmese) (Labov et al., 1972), “almost all dialects of American English” (Labov et al., 1997) and Norwich (U.K.) English (Labov, 1994),28 as well as Scots, South African, New Zealand, and southern British dialects (Lass, 1988). Dressler (1974) identifies numerous other instances of context-free /u/-fronting,29 both with and without associated /o/-fronting. Among the languages he mentions are Oscan, Umbrian, Gallo-Romance, some Iranian dialects, Irish Gaelic dialects, Brithonic, Mingrelian (Caucasian), Xopic (Caucasian), Holoholo (Bantu), Albanian, Frisian, and Parači (Indo-Iranian). Hock (1986) supplies the additional example of Slavic [u:] centralizing and losing its rounding.

The high frequency of /u/-fronting of all these types supports the intuition that “there is certainly no general requirement that any particular high vowel slots in a system be filled…. The most that can reasonably be said is that maybe at least one high vowel slot per system ought to be filled, and back isn’t necessarily the one” (Lass, 1988: 399). In contrast to the abundance of [u] > [y] (> [i]) changes, examples of [y] > [u] are vanishingly rare, and Dressler (1974: 96) attributes each of them to the identifiable “influence of a neighboring language or a substratum or superstratum.” Labov was quite

27 For the curious: Pattern 1 shifts involve the diphthongization of high vowels and the raising of the other long vowels, as in the Great English Vowel Shift. Pattern 2 shifts involve the fronting and subsequent raising of a long, low back vowel, plus the falling of the short front vowels.

28 Labov (1994: 230) argues that in the Norwich dialect, original /u/ fronted, /o/ raised to /u/, and then that new /u/ fronted, creating a situation with two front rounded vowels. The Norwich situation is in that respect similar to the Scandinavian system.

29 This includes changes of [u] (> [y]) > [i].
right to attribute special significance to /u/-fronting and /o/-raising, in contradistinction to their opposites.

One serious question remains: what does one make of these changes? Should phonologists call them push chains, pull chains, mixed chains, or something else entirely? There is hardly consensus on this issue. The chain shift of [u] > [y], [o] > [u] – in that order—receives different interpretations by different authors. Bynon (1983), Bubeník (1983), and Ahn (2004) say that this sequence constitutes a pull chain initiated by /u/-fronting, while Hock (1986: 157) invokes both pull and push mechanisms, saying that “the overcrowding which gives rise to u-fronting and to the ensuing drag-chain developments commonly results from the ‘push’ of a new segment into the back vowel system.” Though Eliasson (1992) disagrees with such an analysis for Scandinavian, he implicitly agrees with Hock’s categorization for a hypothetical case. Both these positions seemingly contradict Martinet’s (1952: 133) intent; he states that “the suggested distinction between drag and push would often be blurred” in practice because “we have to deal with pressure everywhere” in cases like the São Miguel shift.

I believe in taking Martinet’s warning seriously. Rather than forcing these shifts into push, pull, or mixed categories, which are radical oversimplifications in /u/-fronting cases, I suggest using Labov’s terminology instead. “Pattern 3” is a useful designation for the shifts we have discussed in Chapters 3 & 4 because it provides descriptive information about the nature and direction of the chains from a holistic perspective, whereas “push,” “pull,” and “mixed” only indicate the possible causation of individual steps. As a corollary of being causation-based, the traditional method of classification introduce subjectivity, which should also be dispreferred; as we seen several times already, linguists often fail to agree about the causes of a particular sound change. Simply designating /u/-fronting shifts as “Pattern 3” has the benefit of neutrality with respect to this issue.

As a point of historical interest, I believe one reason why the issue of push chains has been so controversial over the past fifty years stems from inconsistency in Martinet’s
Martinet first gives an abstract example of a push chain that clearly involves direct encroachment: “a phoneme A drifting in the direction of a phoneme C… C, instead of awaiting the impending merger, recedes before the invader” (1952: 132). When he presents a concrete example, though, he cites the São Miguel case, which does not involve encroachment but instead, “pressure exerted upon /u/ by the other three back phonemes of a series where margins of security are, by nature, narrower than in the corresponding front series” (1952: 133). No one contests Pattern 3 cases like the São Miguel example; the controversy centers around the hypothetical example. Those who take the abstract definition literally seem to think that Martinet thought all push chains involved encroachment, in spite of the example a few paragraphs later. This leads to the mistaken impression that “in order to be a chain shift, on the push reading there has to be a (near) collision. Push chains require that phonemes bump like balloons, billiard-balls in a rack, on a model resembling the Brownian motion of particles” (Stockwell and Minkova, 1997: 286), or that “push chains would give a relative chronology with considerable overlapping between the steps” (Anttila, 1989: 112). Looking at the spirit rather than the letter of Martinet’s introduction to push chains, we see that what he considered push chains are quite common indeed, but that push chains in the strict sense employed by other linguists do not actually exist. This inconsistency makes it all the more appealing to get rid of this nomenclature altogether.

Now that we have discussed three Pattern 3 cases, we will turn to a more recent—and more complicated—chain shift. This chain, the American Northern Cities Shift, is alleged to include a push chain involving encroachment and has a character very distinct from the shifts we have seen in previous chapters. Examination of this shift will provide additional insights and help further refine our conceptions of what types of chain shifts actually occur in linguistic history.
5. Chapter 5 – The Northern Cities

William Labov has made a career out of tracking the various sound changes that have produced the fascinating and complicated dialectal situation in North American English. His first grand survey of the major U.S. dialects (Labov et al., 1972) was published more than thirty years ago. In this work, Labov and his collaborators identified several vocalic chain shifts that were then—and are still—in progress. One of these shifts, called the Northern Cities Shift, is particularly interesting for our purposes, since recent work by another sociolinguist, Penelope Eckert, purports to show that the two most recent links in the Northern Cities Shift form a push chain of a different nature than the ones we explored in Chapters 3 and 4.

Labov credits Ralph Fasold (1969) with discovering the early parts of the Northern Cities Shift (NCS) in his research on Detroit speech. Today linguists know it to be of much larger scope. The NCS is an urban phenomenon that “has been observed in all the major cities in the Northern dialect area, from the White Mountains in Vermont westward: Rochester, Syracuse, Buffalo, Cleveland, Detroit, and Chicago” (Labov, 1994: 185). With no fewer than six individual steps, Labov (1991: 19) calls it “the most complex case of systematic chain shifting among short vowels, and in fact, one of the most complex chain shifts ever recorded.”

5.1 Methodology

Unlike the shifts discussed earlier, the NCS is very apparent in present-day speech; one need not sift through ambiguous orthographic clues in order to ascertain its course of progression. Linguists are fortunate enough to have extensive audio recordings of speakers who manifest the NCS, as well as sophisticated computer software with which to analyze this ample data.

The majority of evidence for the NCS comes from twenty-five interviews conducted by Labov, Benji Wald, Malcah Yaeger, Sharon Ash, and other researchers in
conjunction with the Project on Cross-Dialectal Comprehension (CDC), directed by Labov. The recordings document interviews taken between 1970 to 1991 in three locations: twelve from Detroit, nine from western New York state (Buffalo, Rochester, and Chili), and four from Chicago (Gordon, 2001). Each vowel token in each interview has been painstakingly labeled, measured, plotted on a grid of F1 vs. F2, grouped into phonemic categories, and compared to other speakers’ vowel systems. From this data, we have a working understanding of the first four steps of the NCS. In the interest of brevity, I will not set out to re-establish the chronology of the first three steps, since it is uncontroversial and will have no bearing on our discussion of the less obvious later steps. The interested reader should consult Labov’s work for ample documentation of the early NCS.

The two most recent steps in the NCS happened too late to be immortalized in the early work on the relevant dialects. The available information about these later steps comes from Eckert’s interviews at a suburban Detroit public high school during the late 1970s. She observed one group of students throughout their junior and senior years, making recordings of two hundred students, out of which she selected fifty-two as a sample (Eckert, 1988: 199). At this school, there were two major social groups occupying opposite ends of the spectrum and correlating to some degree with socioeconomic status. “The ‘Jock’ category, comprised primarily of middle-class students who have enjoyed prominence and faculty favor,” included most of the college-bound students, and the ‘Burnout’ category included those who had enjoyed fewer privileges and were “intending to leave secondary school for the blue collar work force” (Eckert, 1988: 188). These groups primarily differentiated themselves from each other by their attitudes towards authority, manners of dressing, and styles of speech. While most of our data on the latter phases of the NCS comes from these Detroit teenagers, recent CDC interviews from Chicago show that these sound changes were also taking place in at least one other NCS locus from the late 1970s through the 1980s (Labov, 1994: 191).
5.2 Completed phases: Labov et al.

The NCS can be schematized as a sequence of six individual sound changes. I list these steps in (5.2.1) below, showing how my preferred numbering differs from Labov’s.

(5.2.1) Northern Cities Shift steps

1. Raising of /æh/ to /ih/ ([æ] to [iə])          1. Diphthongization of /æ/
   bad → b[iə]d                                          
2. Fronting of /o/ ([a]) to /æh/
   lock → l[æ]k                                          2. Fronting of /o/ to /æ/
3. Centralization/fronting of /oh/ ([ɔ]) to /o/
   call → c[ɑ]ll                                          3. Unrounding/lowering of /oh/
4. Lowering of /e/ ([ɛ]) to /æ/
   head → h[æ]d                                          4. Backing of /ʌ/
4’. Lowering of /i/ ([i] to [e])
   bid → b[e]d                                            5. Change of /e/ to /æ/ or /ʌ/
5. Backing of /e/ to /ʌ/
   bed → b[ʌ]d                                            6. Lowering of /i/ to [ɛ]
6. Backing of /ʌ/ to [ɔ] or [o]
   bus → b[ɔ]ss

The first step in the NCS is very noticeable, even to the untrained ear. The vowel in words like man, hat, bath, and bag has tensed and diphthongized to something like [iə] or [iə] in this dialect. Labov calls this change “raising,” but I believe this leaves out

---

30 I put Labov’s idiosyncratic phonetic notation in slashes and IPA equivalents in brackets where they differ. In sources other than his 1994 book, Labov often leaves out step 4’ in his descriptions of the NCS. Additionally, in the 1994 volume, step 2 is listed as “fronting of /a/,” though elsewhere Labov avoids that symbol altogether, instead using /o/ to indicate IPA [ə]. The reader is excused for finding these inconsistencies confusing!

31 Hence Labov’s notation of this vowel as /æh/, to distinguish it from the lax [æ] of other dialects. /æ/ is tense in certain environments in Southern and Mid-Atlantic dialects, but in the Northern dialect, it is tense everywhere and does not contrast with lax /æ/ (Labov et al., 1972).
important information. As Matthew Gordon (2002: 255) points out, Labov provides two-dimensional (F1 vs. F2) graphs that “offer a very limited picture of the phonetic information available in the speech signal. Moreover, the graphs only chart movement of vocalic nuclei, which means that diphthongization is hidden. Without listening to the audio recordings of Labov’s subjects,\textsuperscript{32} one would surmise from the printed material that [æ] become [ij] for NCS-affected speakers. This is particularly noticeable in the graph of Debbie S.’s speech, partially reproduced in (5.2.2) below:

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{vowel_space.png}
\caption{(5.2.2) Upper front quadrant of Debbie S.’s vowel space, from Labov (1994: 190). F1 runs along the y-axis, and F2 runs along the x-axis.}
\end{figure}

Despite appearances to the contrary, speakers like Debbie S. retain a three-way distinction between bead [bijd], bid [bɪd] and bad [bəd]. Michael Kenstowicz (p.c.) hypothesizes that the diphthongization of /æh/ is the primary change, and the raising of the nucleus arises secondarily through a dissimilatory process that commonly affects centering diphthongs. This would be in keeping with Hock’s (1986: 146-147) observation that tense vowels tend to diphthongize. Whatever the exact mechanism that produced it, like all of the other steps in the NCS, this first step is “tightly correlated with the size of the speech community. The larger the city, the more advanced the change” (Labov, 1994: 32).

\textsuperscript{32} These are readily available with the new online edition of Labov, Ash, and Boberg’s \textit{Atlas of North American English}, available for a free trial from publisher Mouton de Gruyter’s website.
It has therefore taken hold particularly strongly in Chicago, the largest city involved in this series of shifts. Further evidence for the early date of /æ/ raising comes from its large geographical distribution. Dialectologists have noted this same change taking place as far west as the Mississippi River and as far east as Boston (Gordon, 2001: 19).

The next step in the NCS involves the fronting of [α] (Labov’s /o/ or /a/), the vowel in got, which moves into the space vacated by /æh/. It was this shift that first piqued Labov’s interest in the NCS, when in 1968 he misperceived a Chicago teen’s pronunciation of “John” as “Jan” (Labov, 1994: 185-186). One wonders if some ancient Greeks experienced confusion similar to Labov’s when they encountered Attic-Ionic speakers whose dialects had undergone the same change! The fronting of [α], like the diphthongization of /æh/, was “well established even in controlled speech” by the time of the first Chicago recordings in the late 1960s (Labov, 1994: 186-187).

By the early 1970’s, young speakers across NCS territory began exhibiting centralization and lowering of [ɔ] (Labov’s /oh/). This vowel has merged with [α] in many dialects in a process known as the cot/caught merger. In the Northern Cities, however, the (pre-NCS) vowel [ɔ] in caught has remained distinct from the (pre-NCS) vowel [α] in cot. The third step of the NCS had the effect of changing old [ɔ] into [α], a space previously occupied by /o/ (Labov, 1994: 183).

Some recordings from the Northern Cities also show that [i] (Labov’s /i/) has lowered into the mid vowel range (Labov, 1994: 187-188). The chronology of this shift will gain importance as we discuss the most recent steps in the NCS, as reported by Eckert, and I will depart from Labov’s analysis at that point. Thus far, everything I have said about the early NCS phases matches Labov’s own conclusions. We now turn to the later phases, described primarily by Eckert and synthesized into Labov’s overall picture of the diachronic NCS situation.
First, I schematize the steps of the NCS we have seen thus far in (5.2.3) below. The numbered steps in this diagram correspond to the numbered steps in the list in (5.2.1).

(5.2.3) The Northern Cities Shift: Part I

5.3 Ongoing phases: Eckert

Eckert (1986; 1988; 1991) reports that, in her research of Detroit high school students, dialectal features varied according to the social self-identification of the student. Her research was primarily concerned with the social factors of sound change, not fitting the observed changes into the larger picture of the NCS. For this reason, she gives little in the way of a chronology, but in his more recent work, Labov (most notably 1994) attempts to place the changes Eckert noticed within the NCS pattern.

The major way in which Burnouts and Jocks differentiated themselves linguistically was by their contrasting pronunciations of the vowel sound in get. The Burnouts exhibited backing of [ɛ] (Labov/Eckert /e/) to [ʌ], while the Jocks instead tended to lower [ɛ] to [æ]. Adam Albright (p.c.) suggests that perhaps the Jocks had more room for [ɛ] to lower because they had not undergone such a pronounced process of [ɑ] fronting to [æ]. This would be consistent with the Jocks being more conservative than the Burnouts, as other measures show them to be.

In addition, Eckert’s research shows that [ʌ] is also undergoing movement in Detroit. Both she and Labov (1994) claim that it is backing, but the data does not support
such an unequivocal conclusion, as Gordon (2001) briefly notes. Eckert (1988: 200) herself writes that “variants of (uh) in the Detroit area range from the conservative [ʌ] to extremely backed [ɔ], and rounded [ʊ], and to an extreme lowered variant [ʌ]…. In addition, there are small numbers of occurrences of the fronted variants [ɛ] and [ɪ].” From such contradictory evidence, one cannot simply pick one variant and claim that this is the true trend. Whatever is going on with /ʌ/ in Detroit, it is definitely complicated, and it is definitely not consistent.

The problem is further compounded by the fact that Eckert gives no quantitative acoustic data and confounds all the “backed” variants of /ʌ/. She uses a numerical recording system “0 for nonbacked or lowered variants, 2 for the extreme variants [ʊ], [ɔ], and [ʌ], and 1 for intermediate variants (backed and lowered [ʌ])” (Eckert, 1988: 200). This matters a great deal, since the three “extreme” variants show three different types of changes: raising plus rounding, simple rounding, and lowering. There is no way of knowing what she treated as the “original” state of things, and in particular what the difference between 0 and 1 might be. Furthermore, there is no mention of how any fronted variants may have been categorized. Finally, Eckert’s system of simply tallying the point values of the tokens uttered by each speaker means that the “extreme” variants, which include raised and rounded vowels, are conflated with the “intermediate” (i.e., backed and/or lowered) variants in the reporting of scores.

Now we have discussed each of the six individual NCS steps. These are summarized in (5.3.1), with numbering corresponding to the earlier figures (5.2.1) and (5.2.3).
(5.3.1) The complete Northern Cities Shift

We will now turn to the ordering of the more recent steps, which Labov considers to be incontrovertible empirical evidence that push chains do indeed exist. I believe that this evidence, as I have already hinted in Figure (5.2.1), can and should be interpreted differently; there are both empirical and theoretical reasons to suggest that the more recent phases of the NCS did not happen in the order in which Eckert and Labov claim they occurred.

5.4 Analysis

Labov (1994: 200) states that “real-time evidence confirms… push chain status for the backing of /e/” and that “the approximation of /e/ to /ʌ/ is a product of the past 20 years, and that the backing of /ʌ/… is even more recent and by no means as general.” There are several reasons why I believe this change was not as recent as Labov claims. First, in the Chicago interviews that supposedly show the same set of changes as Eckert’s Burnouts, /ʌ/ is not backed relative to the tokens from much earlier Chicago interviews. Labov uses the 1976 interview of the Chicagoan Debbie S. (Labov, 1994: 190) as his best example of unchanged, central /ʌ/ (F2 ~1800-1600Hz), but compared with the rest of the Chicago speakers Labov cites, Debbie S.’s /ʌ/ appears anomalous. Tony C. and Carol Muehe, both teenagers when they were recorded in late 1960s, show /ʌ/ with F2 of ~1300-1200Hz (Labov, 1994: 186-187). Not only did these speakers exhibit /ʌ/ substantially more back than Debbie S.’s, not even the 1989 recording of Jackie H.,
which is supposed to be after the backing of /A/ took place, has F2 that low. Rather than providing evidence for the backing of /A/ during the time frame in which this change allegedly took place, the recordings actually show that /A/ was quite back even at the beginning of this interval.

Additionally, if the backing of /e/ truly provided the push that caused /A/ to move, it is very curious that speakers who lower /e/ rather than backing it would still exhibit “backed” /A/ (e.g., Jackie H. from Labov, 1994: 192). This suggests that, whatever happened to /A/, it was not motivated by the backing of /e/. On the other hand, very early recordings from smaller, more conservative cities in the NCS area show centralized /A/. There was, then, some backing of /A/ early in the NCS. This makes sense since /A/ could certainly have begun to shift back just as soon as /oh/ moved towards [ə], leaving /A/ with no neighbor in that direction. I propose therefore that /A/-backing occurred prior to movement of /e/ in Chicago. The situation remains less clear for Detroit as a result of the confounded data, but if we trust Eckert’s observation that Jocks who lower /e/ also exhibit /A/-backing (relative to early recordings from conservative cities), the same argumentation should hold.

Finally, we must fit the lowering of [i] into the chronology. Labov (1994: 187-188) suggests that this shift forms a drag chain with (i.e., came after) the lowering of [ɛ]. However, he places the lowering of [i] prior to the backing of [ɛ]. Gordon (2001: 10) rightly notes that Labov’s explanation of these events “has a teleological ring to it and raises questions about the underlying motivations for chain shifting;” according to Labov, [ɛ] lowers until it begins to overlap with its lower neighbor [ə], and only then does it undergo “reorientation” and begin backing. This not only teleological, despite the fact that Labov claims to eliminate the role of teleology in explaining chain shifts, but also quite implausible. Placing [i]-lowering in this position separates the two changes
involving [ɛ], but there is no evidence that they were temporally distinct. Additionally, I have already argued that the backing of [ɛ] took place earlier than was previously thought. Squeezing [ɪ]-lowering even earlier, and the lowering of [ɛ] even before that, would imply that evidence of all these changes should be apparent over the entire span of recordings from Chicago, since the later step of /ʌ/-backing shows up in recordings from as early as 1968. In only one of Labov’s (1994) examples do we find [ɪ]-lowering (or [ɛ]-lowering) in the Chicago recordings through 1989. It is particularly striking that we see no [ɪ]-lowering even for speakers who have already undergone [ɛ]-backing, which for Labov would entail prior [ɪ]-lowering (see, e.g., Jackie H.’s charts in Labov 1994: 192-193).

In light of all these findings, I again reiterate my revised chronology of the NCS, as presented earlier in Figure (5.2.1). Diphthongization of /æ/, followed by raising of the nucleus to [ɪ] through a process of dissimilation from the centralized offglide, began the chain shift. The position vacated by [æ] was then filled by the fronting of former [ɑ]. Its neighbor [ɔ] followed suit, unrounding and lowering to become new [ɑ]. Next, /ʌ/ backed, filling the now-large gap in the back vowel series. Further developments involving /ʌ/, including raising, rounding, and lowering, remain unclear. Later, some dialects lowered [ɛ] to [æ], while others backed it to [ʌ]. Finally, [ɪ] lowered into the mid front range.

Each of these steps is compatible with an analysis of motion towards maximal perceptual distance, and none exhibit “push chain pathology” of either the overcrowding

---

33 This speaker, Carol Muehe, exhibits the completion of each NCS feature (except for [ɛ]-backing, which is mutually exclusive with [ɛ]-lowering), so any case her vowel system could not be used as evidence for Labov’s position here, since it does not show only ostensibly early phases, to the exclusion of later ones.
or encroachment type. In this case we actually find ourselves in a happier position than in the Greek, Portuguese, and Scandinavian cases, since we can easily explain the first stage in the shift, thereby solving the inception problem for the NCS. In English, tense vowels tend very strongly towards diphthongization—one does not typically find monophthongal [i], [e], [o], or [u]. The dialect group we have been examining tensed [æ] in all positions, as I noted earlier. Thus, the diphthongization of this vowel had strong precedent. The raising of the nucleus in the new diphthong due to dissimilation created a case vide, in Martinet’s terms, and from there, a complicated pull chain arose, redistributing nearly all the vocalic phonemes.

By the time it was [e]’s turn to move, it had no immediate neighbors either below or behind it, so exemplars more extreme than the mean in both these directions had low probabilities of miscategorizations. The seemingly idiosyncratic Jock and Burnout treatments of [e] seem not to be entirely arbitrary, though. Since the Burnout group was less conservative, their dialect had more /ʌ/-backing as well as a more pronounced shift of [ɑ] to [æ], so there was more room to the back of [e]; for the Jocks, on the other hand, /ʌ/ may have been less backed and [ɑ] not as fronted, resulting in more room below [e]. To complete the chain, [ɪ] lowered, shifting into a position slightly higher than the one formerly occupied by [ε].
6. Chapter 6 – General discussion & conclusions

6.1 Generalizations about attested “push chains”

In the previous three chapters, we looked closely at the four examples of chain shifts commonly cited by linguists as evidence that push chains are empirically attested. We discovered that three of these four cases—Attic-Ionic Greek, São Miguel Portuguese, and Swedish & Norwegian—are nearly identical, not only in the overall effect of the chain shift but also with respect to the order in which the individual steps took place. From knowledge of these languages and the similar cases listed in §4.4, we drew several conclusions about the nature and causality of these Pattern 3 shifts in an attempt to solve the “general problem” posed by Coseriu (1958). I also proposed in this section that “Pattern 3” is indeed a better descriptor of these shifts than the traditional push/pull/mixed chain nomenclature, which I argued on the basis that “Pattern 3” is both more accurate when viewing the chain shift as a unified phenomenon and less subjective because it sets the question of causality aside.

Chapter 5 brought a case of a different sort, the ongoing Northern Cities Shift in American English. This complicated, digital-age series of vowel changes presented a new set of challenges, forcing us to evaluate evidence from audio recordings rather than old texts. The phonetic developments of the shift were also quite dissimilar from the ones in the other case studies. While the Pattern 3 shifts all began with /u/-fronting, the NCS began with the diphthongization of [æ] and subsequent raising of its nucleus through polarization. Importantly, we discovered that the later steps in the NCS do not involve any encroachment or any other peculiarities that would indicate anything but a pull chain ordering. This new interpretation of the evidence directly contradicts Labov’s (1994: 200) statement that “real-time evidence confirms… push chain status for the backing of /e/.”

None of the cases in Chapters 3-5 showed the slightest evidence of phonemic encroachment of the sort that is supposed to characterize push chains. As is particularly
evident in the Pattern 3 cases, instead what we repeatedly saw was one phoneme moving away from a crowded space\textsuperscript{34} and its neighbors then spreading out more evenly in the now-less-crowded space. This is precisely the type of action that typically characterizes pull chains, which leads me to the conclusion that push chains, as traditionally defined, do not exist.

If all chain shifting conforms to these newly imposed constraints, as we have every reason to believe, then it must be the case that many previous theories are far too liberal in their predictions. Of all the theories discussed in Chapter 2, only the feature-based model of discrete sound change and the more recent Blevins (2004) approach make predictions that correspond to the attested variety of chain shifts. The rest of the theories are invalid since they predict that push chains can be started by encroachment of one phoneme causing another phoneme to move. This is no accident—the notion of a push chain is predicated upon this idea of encroachment, in theory if not in practice, since chain shifting of this type simply does not occur.

Of course, one way or another, every theory must allow for mergers of phonemes, which entails admitting either gradual encroachment leading to gradual coalescence or immediate featural change causing instantaneous merger. I have not mentioned merger from a theoretical standpoint in this thesis, except to say linguists have struggled with how to predict when mergers occur as opposed to push chains, and that this dilemma spawned much of the research on functional load. Now that we have removed the idea of encroachment from the theory of chain shifting, that problem disappears. There is no critical factor that influences whether an encroaching phoneme will ultimately merge with its neighbor or push its neighbor out of the way, because we now know on the basis

\textsuperscript{34} It bears repeating that this does not mean that overcrowding is the sole reason for the initial change, but it is certainly one important risk factor. This goes nicely with the intuition that “the full chain shift we call pattern 3... takes place only when there are four degrees of height in back” (Labov et al., 1972: 105) though we now have at least one counterexample, also provided by Labov (1994), in the American West.
of the empirical evidence that chain shifts do not develop in this way from encroachments. In this way, we also further our understanding of phonemic merger by constraining the theory of chain shifting in the manner I have proposed.

6.2 How do chain shifts begin?

With a more accurate picture of the empirical data regarding chain shifts in hand, we can now revisit the more theoretical question of what causes the chain shifts we know to occur. This question can be broken up into two parts: how do chain shifts start (Lass’s 1976 “inception problem”) and how do they continue once the initial step has taken place?

As for the inception problem, there is no reason to expect a single answer prima facie; one might expect that any phonemic drift can trigger a chain shift. In the previous section I identified one major constraint, which is that phonemic encroachment does not trigger (push) chain shifting that begins with movement of the encroached-upon phoneme. I leave it for future research to ascertain whether other such constraints exist, but in the meantime, let us say for the sake of argument that whenever one phoneme (call it phoneme A) moves away from another (call it phoneme B), phoneme B may move into the space vacated by phoneme A, regardless of the reason why phoneme A moved. This should be reminiscent of Labov’s (1994: 118) definition of a minimal chain shift: “a change in the position of two phonemes in which one moves away from an original position that is then occupied by the other.” Labov’s definition is essentially compatible with the empirical evidence put forth in this thesis, though both are irreconcilable with some of the interpretations of attested chain shifts that Labov himself espouses.

Notice also that taking such a broad view of chain shifting makes the inception problem rather trivial: chain shifts begin because a phoneme sometimes moves away from its neighbor. One hardly needs a special explanation built directly into the chain shift mechanism in order to determine why this happens. For example, in §3.4 I outlined some factors that predispose languages towards some degree of /u/-fronting: having four
height distinctions among the back vowels, lacking a high central vowel, and having [u] with the inherent articulatory complexity of the [+ATR, +back] combination. These are risk factors not only for developing /u/-fronting, but also for developing the full Pattern 3 shift. The Northern Cities Shift began with diphthongization of [æ], which requires a different explanation, perhaps related to the tendency for tense vowels to diphthongize and to the tendency for nucleus-glide polarization in diphthongs (Hock, 1986).

6.3 What keeps chain shifts going?

While various disparate factors may contribute towards the initial step in a chain shift, depending on the specific nature of that shift, the same does not seem to be true of subsequent steps. All chain shifts follow the same basic pattern: the first step creates a hole that the second step fills, and so on down the line. Thus, all steps subsequent to the first occur under conditions that are in a certain respect equivalent: each involves movement into an adjacent vacant space. One may legitimately ask both why and how such movement occurs. I will now try to answer those questions in turn.

The various theories of why chain shifts proceed are very similar, differing only in the way that the general “spreading out” of phonemes is measured. One common view holds that “chain shifts would further show one aspect of the principle of maximal differentiation. There seems to be a universal tendency for phonological space, as defined by the articulatory possibilities... to be divided evenly among the units so that each has maximal elbow room” (Anttila, 1989: 186). Other approaches can be formulated by simply substituting “acoustic” or “perceptual” for “articulatory” in the previous sentence. Both the Labov (1994) and Blevins (2004) analyses discussed in Chapter 2 emphasize the perceptual component. Out of the Structuralist, functionalist tradition comes Dispersion Theory (most notably in Liljencrants and Lindblom, 1972), which emphasizes the inherent tension between ease of production and maximal acoustic/perceptual differentiation. In recent years this theory has gained popularity since it has been
translated, with some modifications, into Optimality Theoretic terms (e.g., Flemming, 1996).

I discuss Dispersion Theory (DT) in much detail elsewhere (Vaux and Samuels, 2004), so here I will keep the discussion brief. The main thrust of DT is that phonological inventories are organized by two competing principles, both of which have their roots in the Structuralist tradition: minimize production effort and maximize perceptual contrast. Most proponents of DT, particularly those who operate within OT, maintain that these principles apply synchronically. Here we see functional teleology creeping back into the picture. Bert Vaux and I argued in our 2004 paper that this is the wrong way to view vowel system self-organization for a number of reasons, but this is tangential to the current discussion. It is more important simply to note that these principles most certainly can be interpreted as applying on the diachronic axis rather than on the synchronic one.

Over the course of time, a language may become imbalanced due to changes motivated by other factors: /u/-fronting, for instance, may cause such an imbalance. And again, over the course of time, this imbalance may be corrected, such as by a Pattern 3 chain shift. This avoids any need for a teleological explanation, providing the system moves towards a distribution that is more favorable to the intersection of perceptual and articulatory ease at every step along the way. Some studies have shown that more dispersed or easier to pronounce systems are easier to learn than less dispersed or pronounceable systems (ten Bosch, 1991; Joanisse and Seidenberg, 1998). To this we can apply the idea that “survival of the fittest” applies to language just as it applies to biological organisms, in the long tradition of which Blevins’ (2004) Evolutionary Phonology is one recent expression. In Darwinian terms, the system that is easier to learn gains on its competitors over evolutionary time. Delving even deeper into this, one enters the realm of speculation: “we could speculate that evolutionary pressures might have caused the innate learning mechanism to favor grammars that optimize perception, production, and/or stable transmission in certain ways” (Kiparsky, 2004: 3).
Though so much can be said about why chain shifts occur, we are still far from being able to identify the more mechanical aspects of how they occur. In Chapter 2, I introduced the major theories that have appeared in phonological literature since Martinet’s time. Earlier in this chapter, I mentioned that only two of those theories make predictions that match the empirically attested range of chain shifts: the feature-based approach and the Pierrehumbert/Blevins exemplar-based approach. Deciding between these two theories is not as easy as one might imagine, given how different they are. The typical objection to the feature-based, instantaneous approach to phonetic change is that feature-based models cannot account for encroachment (cf. Labov, 1994: 200). This objection becomes irrelevant once one realizes that encroachment leading to a push chain does not occur. If we theorize that merger is also instantaneous (i.e., that it does not involve ever-closer gradual encroachment until finally the merger happens), then we do not need to account for encroachment in our theory of sound change at all, and a feature-based model may suit us just fine. This hypothesis can be tested through careful empirical investigation of the nature of phonemic merger.

On the other hand, we have not seen anything that can rule out the exemplar-based approach, either. By finding external causes for the initial steps in chain shifts, one avoids the necessity of motivating changes like [u] > [y] purely on grounds of perceptual distinctiveness. This circumvents the difficulty that Blevins (2004) has with such changes, because movement towards perceptual equidistance should not cause changes in which [u] fronts all the way to [y] in a system that already contains [i]. Though in recent years there has been a definite trend towards using exemplar-based models in phonology, the last word on this matter is certainly yet to be written; evidence for both points of view continues to pour in from psycholinguistic experiments. The work presented here does not provide strong support for either side. Still, my conclusion that genuine encroachment does not exist in chain shifting does call into question the arguments that have been made on the basis of this evidence for the gradualness of sound change and exemplar-based approaches.
6.4 Looking to the future

Linguists’ understanding of chain shifting has increased dramatically over the fifty years since André Martinet first introduced the concept of the chain shift. I hope that the present study has contributed to that understanding by examining—and, hopefully, resolving—the push chain debate from both empirical and theoretical angles. Still, a great many mysteries remain about this particular type of phonological change. At the outset of this thesis, I mentioned five problems that apply generally to the study of language change: the rational, general, historical, actuation, and inception problems. I noted on the first page of the introduction that “push chains are interesting from the perspectives of all these problems: there are no circumstances under which they necessarily occur, the circumstances under which they may occur are far from clear, the mechanism by which they occur has not been explained, and the reasons why they occur seem murky.” We can put these problems to rest as far as push chains are concerned simply by saying that they do not exist; the case studies of putative push chains in Chapters 3-5 fully support this conclusion.

Some problems still remain, however, when we look at the kinds of chain shifts that do exist. We still have not identified conditions that are both necessary and sufficient to trigger chain shifting, though in historical linguistics few things are inevitable, and such surety is almost certainly too much to ask. We have, however, made progress towards determining conditions under which Pattern 3 shifts, at least, may occur. We were able to narrow down the list of possible mechanisms for chain shifting out of the group that has been proposed, but we could not decide on just one in particular, or even between the classes of feature-based and exemplar-based models. By eliminating push chains from the picture, we have simplified the picture of why (pull) chain shifts happen—a phoneme moves, leaving an open space—but beyond this general principle, details are lacking. These are not intractable problems, but they are demanding ones.
In short, much work remains to be done on chain shifting, and I sincerely hope that historical linguists and phonological theorists alike will devote more time to this endeavor. I will conclude the present treatise by giving some ideas for future research that I believe could help shed light on the unresolved matters. For instance, now that powerful computers are readily accessible, phonologists can move beyond the era of two-dimensional graphing of acoustic space. Coupled with the ease and accuracy of digital recording equipment, computer-aided modeling of acoustic measurements would doubtless produce some surprising results. Longitudinal studies tracking speech communities over many years would make excellent data sources for such analyses. The study of phonological acquisition, both by children and by adult language learners, is a field that continues to grow, and as it does so, our understanding of diachronic phonology will increase correspondingly due to the close relationship between the acquisition process and language change. Perceptual studies aiming at determining what speech sound contrasts can and are perceived would also further our understanding of how the learning mechanism is constrained, and how phonological systems are constrained as a result. Similarly, we would be well served by further inquiries into the nature of the mental statistics that people keep about language, because an increasing number of theories, such as Pierrehumbert’s (2001), are dependent on particular statistics being kept. I alluded to this matter in a footnote in §2.7. Last but certainly not least, we should also strive to document as many of the world’s languages as possible. Generalizations are only as good as the data over which the generalization is made, so if the data is skewed by our ignorance, our theories will almost certainly be wrong. These projects should keep many future generations of linguists busy!
7. References


—. 1926. Sognemålene. Oslo.
MATHESIUS, VILÈM. 1929. La structure phonologique du lexique du tchèque moderne. Travaux du Cercle Linguistique de Prague, 1.67-84.
MCGINN, RICHARD. 2005. What the Rawas dialect reveals about the linguistic history of Rejang. Oceanic Linguistics, 44.12-64.


ROGERS, FRANCIS MILLET. 1940. The pronunciation of the Madeira and Azores dialects as compared with standard Portuguese. Harvard University: Ph.D. dissertation.


