Chapter 2

What is Phonetic Change?

In this chapter, I will lay out the most basic description of the phenomena I will be addressing in this dissertation, provide some necessary terminological clarification, outline my minimal theoretical commitments in carrying out this project, and most importantly, highlight where my results and analysis diverge from previous work on the topic, and why they are of interest to phoneticians, phonologists, and sociolinguists.

2.1   Sound Change and Grammar

I will be using the term *sound change* to cover a broad range of phenomena, including phonemic mergers, lexical diffusion, Neogrammarian sound change, rule loss, rule generalization, etc., and I will use the term *sound system* to broadly refer to the domain of language where sound change takes place. I will be reserving the terms *phonological change* and *phonetic change* to refer only to changes which occur within the domain of phonology and phonetics, respectively. To the degree that any particular sound change is ambiguous between whether it takes place within the phonological or phonetic domain of language, it will be ambiguous as to whether these changes should be called phonological or phonetic.

Clearly, the the potential for phonology-phonetics ambiguity is vast, and contentious. Pierre-humbert (1990) described many researchers involved in debate over whether phenomena should be described as phonological or phonetic as "intellectual imperialists," and Scobbie (2005) labeled

these debates similarly as "border disputes." However, for the study of sound change, resolving these disputes is not merely a terminological issue. Starting with Labov (1969), it has been established that the structure and formal properties of the grammar one posits makes clear predictions about how the linguistic variation we observe should be structured. In this landmark study, Labov addressed the topic of copula absence in African American English. First, by establishing that copula absence was prohibited under certain structural conditions, Labov concluded that AAE made productive use of the copula (unlike, for example, Russian), thus copula absence must be the product of a deletion process. Then, through a quantitative analysis of the proportions of copula deletion, Labov was able to conclude that copula contraction and deletion were separate processes, and that contraction was ordered before deletion. This early case study highlights the importance of having an adequate grammatical model in order to structure the quantitative analysis of variation. The results of the quantitative analysis can then further narrow the grammatical possibilities.

While Labov (1969) was a purely synchronic case study, the pattern of mutual reenforcement between grammatical theory and language change has also been well established. For example observation of the Constant Rate Effect in syntactic change led Kroch (1989, 1994) to conclude that the locus of syntactic change is within the features of syntactic functional heads. Kroch (1989) was specifically arguing against the "Wave Model" of language change put forward by Bailey (1973), in which it is suggested that those contexts which are most advanced in the direction of language change are i.) where the change began and ii.) moving the fastest. Kroch (1989, 1994) found that for several examples of syntactic change, this pattern did not hold, indicating that the objects of syntactic change in these cases were functional heads, rather than larger collocations, or constructions. Fruehwald et al. (forthcoming) relied on the same analytic technique to argue that the locus of phonological changes are generalized rules which operate over all segments which meet the appropriate structural description. Of course, not all analyses of language change have supported generative-like theories of grammar. Notably Phillips (1984, 1999, 2006) and others have focused on the effect of lexical frequency on the propagation of sound change in support of a usage based model of phonological knowledge. Despite the potentially radically different theo-

retical commitments of the researchers involved, they all share the same analytic commitments: grammatical theory constrains the set of predicted language changes, thus observed patterns of language change serve as crucial evidence for or against one's grammatical theory.

Just as the structure of grammatical theories can be confirmed or falsified through the study of sound change, so can their scope. It is a well established theoretical position that language change and variation necessarily occurs within the non-arbitrary and explicitly acquired domain of linguistic knowledge. For example, Kiparsky (1965) notes that the generative view of language change is that it takes place in the Saussurian *langue*, or generativist competence, rather than in the *parole* or performance. The variationist paradigm has also placed patterns of variation squarely within the linguistic competence of speakers, as Weinreich et al. (1968, p. 125) stated:

> deviations from a homogeneous system are not all errorlike vagaries of performance, but are to a high degree coded and part of a realistic description of the competence of a member of a speech community.

Hale (2004) makes this position very explicit in his chapter on Neogrammarian sound change, where he describes all changes as abrupt disjunctions between the grammar of a language acquirer and the grammar of the speaker who served as their primary linguistic model. From these explicit formulations of sound change as grammatical change follows the conclusion that the structure of one's grammatical theory places a hard boundary on the extent of possible sound changes. Only those aspects of language which are learnable and representable in speakers' knowledge may be subject to change. Now, Kiparsky (1965) explicitly excluded phonetics from grammatical competence, treating all sound changes as phonological. The exclusion of phonetics from linguistic competence has since been relaxed by almost all researchers since some seminal work in the 1980's (Liberman and Pierrehumbert, 1984; Keating, 1985, 1988, 1990) with some notable exceptions (e.g. Hale et al. (2007); Hale and Reiss (2008)), and as I will illustrate in §2.3, the existence of truly phonetic (i.e. continuous) sound change demands the inclusion of phonetics within linguistic competence.

The structure of one's grammatical theory also places a hard boundary on the range of possible typological variation between languages and dialects. A biconditional relationship between

typological variation and sound change therefore follows. For any given dimension of typological variability, there may be a sound change along that dimension, and for any given sound change along a given dimension, there may be typological variation. For example, two languages could conceivably differ in whether or not voicing is contrastive at all points of articulation in their stop series, such that Language A contrasts /k, g/ while Language B has only /k/. The existence of such a typological contrast implies the possibility of a sound change which alters the knowledge of this contrast, merging /k, g/ > /k/ in Language A. Conversely, if we were to observe a phonetic change within one language whereby the duration of the vowel /i/ decreased by 50ms (not contingent on any other sound changes, per the concerns of Hale et al. (2007)), that would imply the possibility of cross-linguistic differences in the duration of vowels, minimally of 50ms, thus the ability of speakers' linguistic competence to represent and control such a difference (Labov and Baranowski, 2006). In the following subsections I review some examples of this biconditional relationship between typological variation and possible sound changes. This is, to be sure, an incomplete list, but is intended to cover perhaps the most common kinds of sound changes and typological differences, with the goal of localizing them to a specific domain of speakers' knowledge.

2.1.1 Phonemic Incidence

One obvious point of cross-dialectal variation is what I'll broadly call phonemic incidence, relating to the phonological content of lexical items. For the purpose of this discussion, this knowledge includes the phonological content and identity of segments within a given lexical item, and their linear order. The Atlas of North American English reports on such an example of cross-dialectal variation in phonemic incidence for the lexical item *on* (p. 189, Map 14.20). Looking exclusively at speakers who maintain a distinction in their low-back vowels between a short, lax, low-back vowel (as in the name *Don*) and a long, tense, low-back vowel (as in the name *Dawn*), Northern speakers place the lexical item *on* in the same phonemic class as *Don*, while Midland and Southern speakers place it in the same class as *Dawn*. Coye (2010) finds the same North-South split within the state of New Jersey, as well as a split according to the first vowel in *chocolate*, which Northern

Speakers classify with the long, tense vowel and Southern speakers classify with the short-lax vowel. These reported facts from the ANAE and Coye (2010) are summarized in (2.1–2.2).

(2.1) North

    ɑ *Don, on*
    ɔː *Dawn, chocolate*

(2.2) South

    ɑ *Don, chocolate*
    ɔː *Dawn, on*

These differences in phonemic incidence between the two dialects cannot be explained in terms of phonological constraints of any sort. Both the Northern and Southern regions allow both /ɑn/ and /ɔːn/ sequences as evidenced by the difference between *Don* and *Dawn*, and there is similarly neither dialect has a constraint against /ɑk/ or /ɔːk/ sequences (.e.g. *tick-tok* [ŋ] and *talk* [ɔː]). Instead, these cross-dialectal differences are due to the arbitrary knowledge about the lexical entries for *on* and *chocolate*.

And just as phonemic incidence can vary cross-dialectally, it can also be subject to language change. Specifically, many cases of lexical diffusion can be described in terms of shifting phonemic incidence, as can phonemic mergers by transfer (Herold, 1990). An example of change easily relatable to the distribution of /ɑ/ and /ɔː/ would be the development of diatonic pairs in English, as discussed by Phillips (2006, Chapter 2, p. 35). Phillips specifically investigates diatonic pairs (minimal pairs of nouns and verbs which differ only in the placement of stress, e.g. *récord*.n ∼ *recórd*.v), where the stress for both parts of speech was originally final. For example, both the verbal and the nominal forms of *address* originally had final stress, but the nominal form has now has initial stress. Phillips (2006) found that of all of the potentially diatonic word pairs, the ones which actually underwent a stress shift from final to penultimate were lower in frequency than those where the stress remained final. Given minimal pairs like *áddress* and *addréss*, the stress placement in these words must be part of their lexical entry. The sporadic, lexically diffuse, and frequency sensitive nature of the change from final to penultimate stress for these words suggests that the locus of this change is in the lexical entries, meaning that the development of

every diatone pair is a separate change of the form *addréss>áddress*. Explaining the fact that there appears to be a systematic and unidirectional development of final to penultimate stress for these lexical items is beyond the scope of this discussion.

I would also classify the presence or absence of phonological material in a lexical entry under the umbrella of "phonemic incidence." For example, Bybee (2007, [1976]) reports on lexically sporadic schwa deletion in English, which she argues is primarily driven by lexical frequency, producing pairs like *memory* and *mammary*, the first being more frequent in use, and more frequent in schwa deletion. How the difference between *memory* and *mammary* ought to be captured depends in part on your theoretical commitments regarding the content of lexical entries. Bybee's own analysis is that [ə] is represented as phonetically gradient in the underlying representation, an analysis which I myself do not adhere to. For the the sake of exposition, I'll suggest that *memory*, for many speakers much of the time, has the underlying representation /mɛmri/, while *mammary*, for most speakers most of the time, has the underlying representation /mæməri/. Guy (2007) also appeals to variable lexical entries in order to account for the exceptionally high rate of TD Deletion for the word *and*. *And* undergoes TD Deletion at a much higher rate than would be expected given other predictors, so Guy (2007) suggests that some proportion of the missing /d/'s is due to their absence in the lexical entry for *and*, meaning there are two competing lexical entries: [ænd] and [æn].

I also include the linear order of phonological content as falling under this domain of knowledge. A very salient example of cross-dialectal variation in the linear order of phonological material in North America is the difference in *ask* between most White dialects (/æsk/) and African American English (/æks/). This is clearly a difference in lexical knowledge rather than, say, the reflex of different phonotactics, because the difference in /sk/~/ks/ order is restricted to only this lexical item. Similarly, there are examples of lexically sporadic metathesis changes. *The Metathesis Website* (Hume, 2000) provides the example of *chipotle* (an increasingly common word in North America due to the restaurant chain named after the smoke-dried japepeño), which is sporadically metathesized /tʃɪpotle > tʃɪpolte/. There are, of course, many more examples of metathesis in sound change, such as those given by Blevins and Garrett (2004). However, Blevins and Garrett
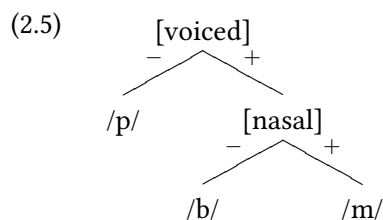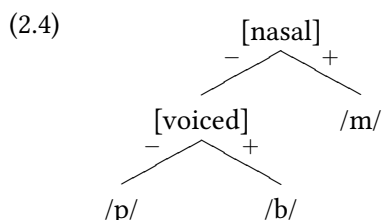
(2004) describe most of their examples of metathesis as fully regular in their outcomes, making it ambiguous as to whether these sound changes progressed as a series of lexically sporadic metatheses, ultimately concluding by spreading across the entire lexicon, or as a the result of a new phonological process or phonotactic being introduced into the grammar elsewhere. This latter option conceptually possible due to productive metathesis processes in synchronic phonological grammars, as Buckley (2011) discusses extensively. Mohanan (1992) and Anttila et al. (2008), for example, describe the following productive alternation for some speakers of Singaporean English.

(2.3)

	Word Final		Intervocalic	
lisp	[li**ps**]	*lisping*	[li**sp**iŋ]
crips	[kri**ps**]	*crispy*	[kri**sp**i]
grasp	[grɑː**ps**]	*grasping*	[grɑː**sp**iŋ]

The locus of this variation is almost certainly not in the lexical entries for these words, but rather in the phonological processes of Singaporean English, a domain of knowledge which I will address in §2.1.3.

2.1.2 Systems of Phonological Contrast

It has been suggested that speakers' knowledge of their phonology includes a structured representation of phonological contrast (Hall, 2007; Dresher, 2009). According to this hypothesis, two languages could differ crucially in the representation of their bilabial stop series in the following way (from Dresher (2009)).

(2.4)

```
           [nasal]
        −  ⁄    ⁀  +
   [voiced]        /m/
  − ⁄   ⁀ +
 /p/     /b/
```

(2.5)

```
          [voiced]
       −  ⁄   ⁀  +
    /p/        [nasal]
            − ⁄   ⁀ +
           /b/     /m/
```

Under the *Contrastivist Hypothesis*, in the language with the contrastive hierarchy in (2.4), /m/ would not participate in, say, voicing assimilation processes, because it is not contrastively specified [+voice]. In a language with the same exact phonemic inventory, but the contrastive hierarchy in (2.5), /m/ *would* participate in voicing assimilation processes.

Recent works supporting the hypothesis that speakers represent contrastive hierarchies such

as those in (2.4, 2.5) have actually turned to patterns in historical language change for evidence. Dresher et al. (2012) and Oxford (2012) cite examples of phonological changes in Algonquian languages, Manchu, and Ob-Ugric languages which appear to involve the demotion of contrastive features down the hierarchy, resulting ultimately in phonemic mergers.

Merger, of course, is one of the most well studied kinds of sound change. Hoenigswald (1960, chapter 8) called merger "the central process in sound change." However, recent studies of merger in-progress (e.g. Herold (1990); Johnson (2007)) have focused most closely on the mechanisms of merger of just two segments, without necessarily discussing the larger effect of these mergers on the larger systems of contrasts in the language. Very recent work in Columbus (Durian, 2012), New York City (Becker, 2010; Becker and Wong, 2010), and Philadelphia (Labov et al., 2013), however, hint that there may be some more systemic consequences of merger, specifically the low-back merger. All of three of these large urban centers exhibited a so-called "split-short-a" system, whereby there was an opposition between a short, lax /æ/ and a long, tense, ingliding version, varying in its phonetics between [æː] and [iə]. I will refer to the long, tense variant as /æː/. In all three locations, the distribution of /æ/ and /æː/ was semi-regular, but in all cases complex, and exhibiting some lexical irregularity. Durian (2012), Becker and Wong (2010) and Labov et al. (2013) all report this complex opposition of /æ/ and /æː/ breaking down in favor of a simple nasal-short-a system, whereby the distribution of /æ/ and /æː/ is totally predictable based on whether or not the following segment is nasal. Another concurrent change in all three of these cities is the lowering of the long, tense, ingliding *back* vowel, /ɔː/, towards the short, lax vowel, /ɑ/ (Durian, 2012; Becker, 2010; Labov et al., 2013). This lowering of /ɔː/ has been followed by the low-back merger in Columbus, and no study of merger in Philadelphia or New York City has been carried out. If (and this is debatable) we were to treat the opposition of /æ/ and /æː/ as being contrastive, we could conceive of the following contrastive hierarchy.

(2.6)

[low]
− +
... [back]
− +
[tense] [tense]
− + − +
/æ/ /æː/ /ɑ/ /ɔː/
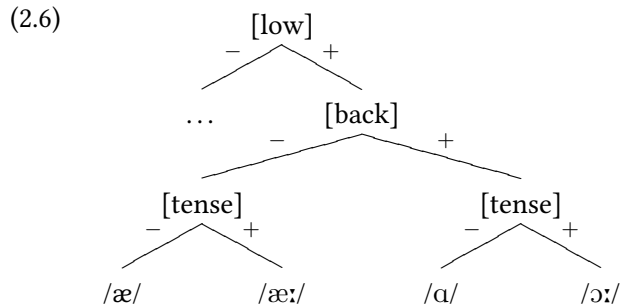
The transition from a split-short-a system to a nasal-short-a system would amount to losing the contrastive [±tense] specification for /æ/∼/æh/ in favor of a purely allophonic distribution. The same would go for the low-back merger. This is, of course, merely a suggestion, intended as an illustration of how attempting to properly localize a particular sound change to a particular domain of linguistic knowledge can serve to both unify seemingly disparate events, and open the door to new and interesting lines of research.

2.1.3 Presence or Absence of Phonological Processes

Perhaps the most discussed cross-linguistic/dialectal difference is the presence or absence of a given phonological process. In serial rule based approaches to phonological theory, this could be captured by the presence or absence of a phonological rule, and in constraint based grammars, by the high or low rankedness of the constraint(s) motivating the process. A good example would be word final devoicing of obstruents, which is a broadly attested process cross-linguistically.

Phonological systems of languages change, logically necessitating the addition or loss of phonological processes to be a possible language change. While most accounts of change of this sort focus on the phonologization of phonetic processes as the addition, and the morphologization of a phonological process as the loss, Fruehwald et al. (forthcoming) and Gress-Wright (2010) examined a case of the loss of a phonological process which appeared to be directly lost without becoming morphologized. In Early New High German (ENHG), there was a productive process of word final devoicing. The relevant alternation is illustrated in (2.7-2.8).

(2.7) "day": [k]∼[g]

(a) *tac* (acc.sg)

(b) *tage* (acc.pl)

(2.8) "strong": [k]∼[k]

(a) *stark* (uninflected)

(b) *starkes* (neut.nom.sg.)

ENHG underwent a process of apocope, which applied variably (at least as determined by the orthographic trends), and produced opaquely voiced word final obstruents.

(2.9) "day": [k]∼[g]

(a) *tac* (acc.sg)

(b) *tage*∼*tag* (acc.pl)

Many dialects of ENHG subsequently lost the process of word final devoicing. Gress-Wright (2010) argues that this was triggered by the opacity created by apocope. It was clearly not a case of general sound change, because it only affected word final voiceless obstruents which were underlyingly voiced.

(2.10) (a) tac; tage > tag; tag

(b) stark; starkes > stark; starkes

The process was also lost as a whole, rather than segment by segment, as Fruehwald et al. (forthcoming) found that the Constant Rate Effect (Kroch, 1989) applied in this case. We compared the rate of the loss of word final devoicing across the voiced stop series (/b, d, g/) and found that the rate of change was the same across all three stops in multiple dialects. We took the presence of Constant Rate Effect in this case as evidence for there being just one phonological process in the grammar which applied all relevant segments. This one phonological process was then gradually lost, affecting all relevant segments at the same rate. Even though the loss of word final devoicing was more advanced for some segments than others, the fact that they all lost the process at a constant rate suggests that the differences in their rates of devoicing were due to properties of language use, rather than differential treatment by the grammar.

2.1.4 Targets of Phonetic Implementation

I won't spend an undue amount of space here discussing how the targets of phonetic implementation can vary cross-linguistically, and change over time, since this is the broad focus of this dissertation. Needless to say, languages and dialects can vary greatly in terms of the phonetic realization of segments which can be considered phonologically identical. An early approach to looking at this was Disner (1978), who compared the vowel systems of various Germanic languages and found that there were not universal targets for vowels which were putatively the same between them. An extreme case is Danish, which has six of its seven vowels in the high to high-mid range, and its seventh as low-central.

A more certain example of phonologically equivalent vowels which differ only in their phonetic realization can be found in /ow/ in North America. Figure 2.1 displays a map from the Atlas of North American English (Labov et al., 2006) which denotes the Southeastern Super Region as defined by the fronting of /ow/. Speakers represented on the map with light red points have /ow/ fronted past the threshold by which the ANAE diagnosed /ow/ fronting. There is no compelling dialectal data to suggest that /ow/ should have a different phonological status in the Southeastern Super Region as distinct from the rest of North America. The largest phonological differentiator in North America is the low-back merger of *cot* and *caught*, and as can be seen in Figure 2.1, the regions with the merger only partially overlap with regions with fully back /ow/.

/ow/ fronting has also been a change in progress in Philadelphia as reported in 1970s (Labov, 2001), but has begun backing (Labov et al., 2013). Figure 2.2 displays the diachronic trajectory of /ow/ along F2, subdivided by men and women and level of education. From the turn of the century until just after 1950, /ow/ fronted dramatically for women who did not go on to higher education. Men, and both men and women with some higher education, participated minimally in this change.

The specific targets of phonetic implementation for the same phonological objects can thus vary cross-dialectally and across social groups, meaning that speakers must be able to represent differences in phonetic targets at least as small as the increment of change for women. §2.3.1 will be devoted to arguing that this incrementation is effectively infinitely small since the change
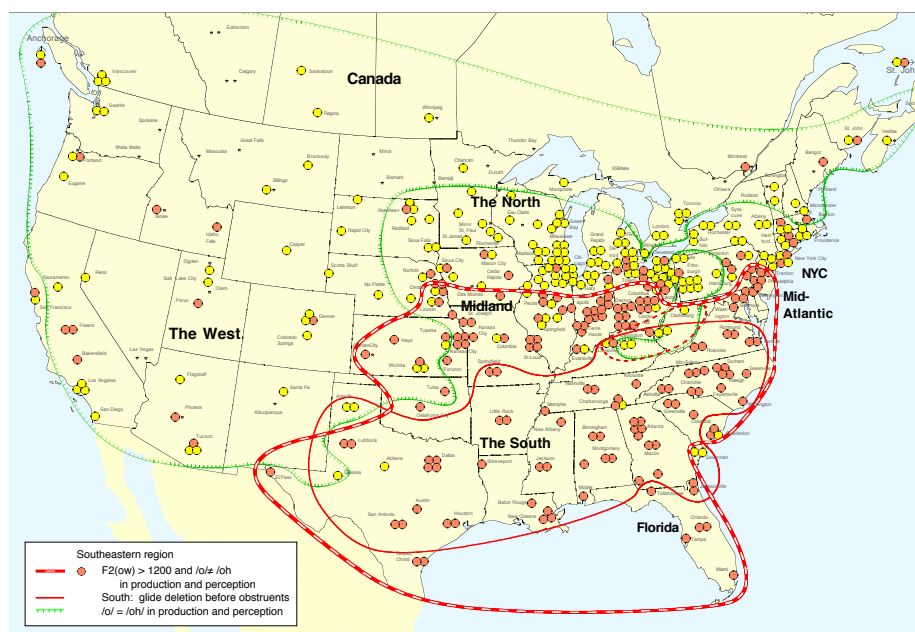
Figure 2.1: The division of North America into /ow/ fronting vs. /ow/ backing regions. From the Atlas of North American English.
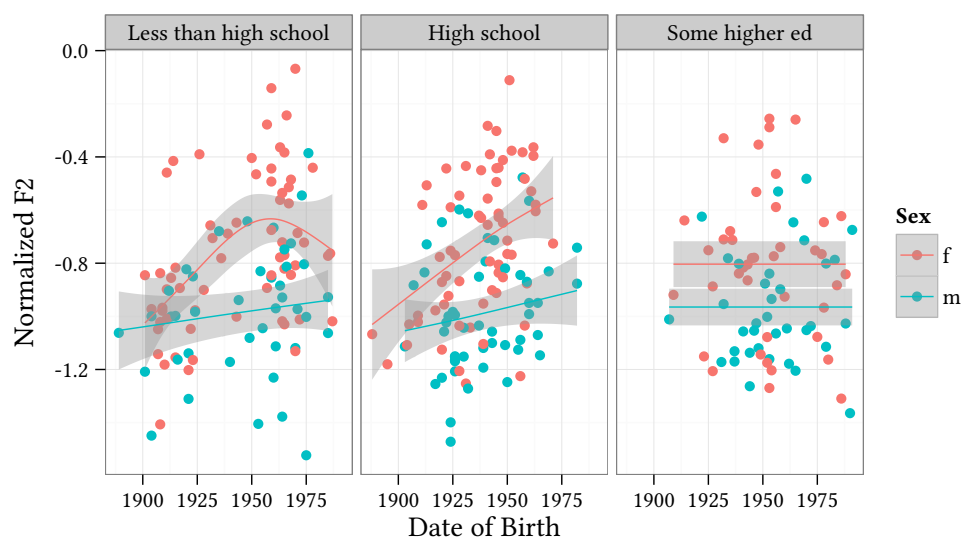

Figure 2.2: The fronting and and subsequent backing of /ow/ in Philadelphia.

is truly continuous, meaning that the phonetic representation must be of a different type than categorical phonological representation.

2.1.5   Gestural Phasing and Interpolation

In addition to the language specific targets of phonetic implementation, there are also appears to be language specific processes of phonetic interpolation and gestural phasing. However, it is necessary to be careful about making such claims, because apparent differences in phonetic interpolation may actually be related to higher level facts, like contrastivity. For example, Cohn (1993) finds that English allows for gradient nasalization of pre-nasal vowels, while French does not. This may at first appear to be a language specific difference in, say, the gestural phasing of velum lowering, but it seems more likely to be related to the fact that French has contrastive nasal vowels, and English doesn't. Oral French vowels have an explicit oral target, while English vowels are allowed to be non-contrastively nasalized.

However, dialectal differences in stop epenthesis in English as reported in Fourakis and Port (1986) seems to be a clearer case of differences in gestural phasing. Fourakis and Port (1986) first argue that stop epenthesis in American English, rendering words like *dense* and *dents* roughly homophonous, is not a phonological process because they found reliable phonetic differences between epenthesized [t] and underlying [t]. Instead, they argued that it results from gestural overlap of the closure from the [n] and the voicelessness of the [s]. However, South African English does not exhibit stop epenthesis in [ns] sequences. If anything, their representative spectrograms of South African speakers seem to show a very brief vocalic period of just 2 or 3 glottal pulses between the offset of the [n] and the onset of the [s]. This appears to be a good case of cross-dialectal variation in phonetic alignment.

A very striking example of language change involving shifting phasing relations comes from Andalusian Spanish. As with many dialects of Spanish, /s/ aspirates in many positions, including before stops, and in Andalusian Spanish, this is also frequently associated with post-aspiration (Torriera, 2007; Parrell, 2012; Ruch, 2012).

(2.11)   *pasta*

pasta > pahta $\sim$ paht^{h}a $\sim$ pat^{h}a

Ruch (2012) found that the duration of pre-aspiration is decreasing in apparent time, and the duration of post-aspiration is increasing in apparent time. Torreira (2006); Torriera (2007) and Parrell (2012) analyze this change in terms of a change in alignment of the stop closure gesture and the spread glottis gesture. If this analysis is correct, then it is a striking example of of a language change affecting phonetic alignment/phasing.

I should note that a model of coarticulation based on phonetic interpolation through unspecified domains (Keating, 1988; Cohn, 1993) versus one based on articulatory gestures and their phasing (Browman and Goldstein, 1986; Zsiga, 2000) propose vastly different mechanics of coarticulation, but for the purposes of this dissertation, their mechanical differences are not of as much consequence as the resulting phenomenon, which is roughly equivalent.

2.1.6 Sound Change and Grammar Summary

The over-arching goal of this section has been to highlight the crucial but non-trivial connection between observed sound changes and the proposed grammars in which they are occurring. I say "non-trivial" because it does not appear to be the case that the domain of knowledge of a given sound change can be determined simply from the outcomes of the sound change. For example, both "merger" and "metathesis" were the outcome of three different kinds of changes in speakers' knowledge.

(2.12) sources of merger

    (a) lexically gradual change in phonemic incidence

    (b) change in the system of contrast

    (c) phonetically gradual change

(2.13) sources metathesis

    (a) lexically gradual change in phonemic incidence

    (b) introduction of a productive phonological process, which outputs metathesis

(c) gradual change in the alignment of articulatory gestures

I should add that this is not meant to be an exhaustive list of all the ways in which merger and metathesis come about, since these are not the focus of this dissertation. Rather, I hope to have made clear that for any given start and end points of a language change, there is not necessarily a unity of process that produced the change. There are many paths between two diachronic stages of a language, both in principle, and attested in the study of sound change.

I also hope to have made clear exactly the role I see the study of sound change playing in the general linguistic enterprise of delimiting the possible knowledge of speakers, rather than simply being a case of "butterfly collecting." The same goes for the high-volume-data and statistical analysis which form the empirical base of this dissertation.

2.2 The Phonology-Phonetics Interface

Having discussed the importance of localizing particular sound changes to specific domains of knowledge, I'll now outline the architecture of the Phonology-Phonetics interface that I'll be assuming in this dissertation.

2.2.1 Modular and Feedforward

In this dissertation, I will be operating within the paradigm of phonology and phonetics which is modular and feedforward, to use the terminology from Pierrehumbert (2006). My motivation for explicitly committing to a particular framework is not, primarily, to argue for the correctness of that framework. Rather, it is in acknowledgement that in linguistics, as with all other fields of scientific inquiry, it is only possible to make progress if we commit to a particular paradigm while performing our investigations. It is the theoretical framework which delineates the set of facts to be explained, and defines how new results ought to be understood. To the extent that a theoretical framework is successful at discovering new facts to be explained, and for pursuing analyses of these facts, we can call it successful. There is thus a mutually reenforcing relationship between the results of research, which would be impossible to arrive at without presupposing

a theoretical framework, and the theoretical framework, which is supported by its results. The same is true of this dissertation.

2.2.2 The Architecture

The grammatical architecture I'll be broadly adopting is that proposed by "Generative Phonetics" (Keating, 1985, 1990; Pierrehumbert, 1990; Cohn, 1993 *inter alia*). The most important, core aspects of this model are the modular separation of phonology and phonetics, and the translation of phonological representations into phonetic representations by a phonology-phonetics interface. A schematic representation of this grammatical system, as adapted from Keating (1990), is given in Figure 2.3.

phonological input

| Phonological Grammar |

surface phonological representation

| Phonology-Phonetics Interface |

phonetic representation

| Alignment & Interpolation |

gestural score/intention

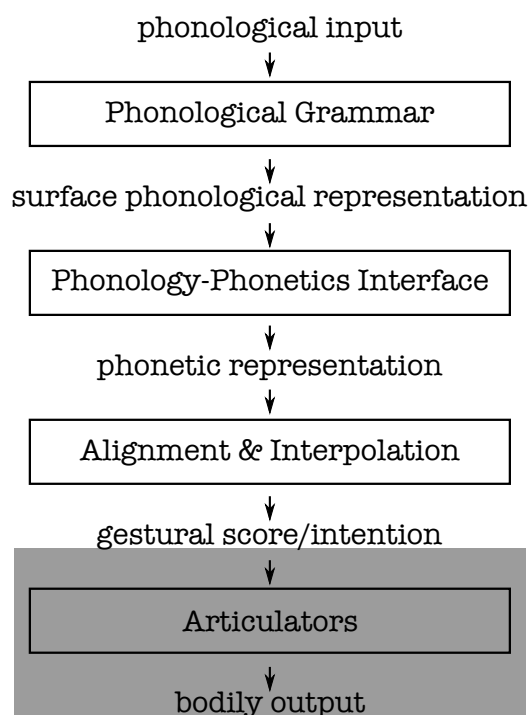| Articulators |

bodily output

Figure 2.3: Schematic of the phonology & phonetics grammar.

In the following subsections, I'll relate each level of the grammar to the discussion above regarding the strict relationship between cross-linguistic typology and sound change, with the understanding that my strongest theoretical commitments in this dissertation regard the Phonology-Phonetics Interface.

Input

The input to phonological computation is underlying form stored in the speaker's lexicon. It is at this level of representation the differences in phonemic incidence, as described above, occur. For the purpose of this dissertation, I have almost no commitments to the nature of this underlying form, such as whether it should be underspecified, and to what extent or based what principles, or what constraints may or may not exist on possible underlying forms. My only theoretical commitment is that the underlying form should be categorically represented. There may be some variation at this level representation, but that would be represented as having multiple possible underlying forms to choose from for a given lexical entry. For example, speakers of African American English may variably choose /æsκ/ or /æks/ for the lexical entry for Asκ. This variation in the *choice* speakers make between underlying forms does not mean that the options themselves are gradient.

Phonological Processing

My assumption about phonological processing is minimally that it maps phonological inputs to outputs which have the *same* representational system. Whether this mapping is done in rule-based serialist framework or a constraints based framework is not of particular importance here. For the most part, I will be describing phonological processes using a rule based notation, but I am not taking that to be a substantive point. I will, however, make some allusions to a layered or stratal model of phonology. This is partially because some of phonological processes I identify apply at different morphosyntactic domains, a fact which is more easily captured by phonological theories which include strata. It is also partially due to the explicitness of the relationship between diachronic change and strata that Bermúdez-Otero (2007) makes.

Targets of Phonetic Implementation

The aspect of the the grammatical architecture in which I have the most at stake is the interface between phonology and phonetics. My initial assumptions about the interface are

(2.14) that it operates over the surface phonological representation,

(2.15) more specifically, it operates over phonological features.

The implication of (2.14, operation over surface forms) is that neither the underlying form, nor the phonological processes which applied to it to produce the surface phonological form can be relevant to phonetic implementation unless their properties are somehow carried forward to the surface phonological representation. This means that when we see two phonetic forms that we have some reason to believe have different surface phonological representations (e.g. low [ɑɪ] and raised [ʌi] before voiceless consonants), we can't determine whether this is because a phonological process differentiates these variants, or whether this is an underlying contrast present in the input to phonological processing without appeal to independent facts.

The implication of (2.15, operation over features) is that surface representations which share phonological features must also share some common phonetic target. This point may appear to be too pedantic to mention, since most feature theories explicitly name phonological features after their phonetic properties (e.g. [±high], [±ATR]), so it would necessarily follow that we wouldn't posit a segment as possessing a feature if it didn't also possess the phonetic property. If we only posit the feature [+ATR] for segments which have the property of advanced tongue root, then it is vacuously true that all segments with [+ATR] will have a phonetic target of advanced tongue root. However, the specific case of phonetic change in combination with some recent rethinking of phonological representation does require assumption (2.15) to be made explicit. I am positing that phonetic change involves changes to phonetic implementation, so that at one time point [+back] has one phonetic target, and at a later time point [+back] has a different target. The question immediately arises as to which vowels ought to be affected by this change, and the answer, given (2.15), is all of those which share the feature [+back].

Furthermore, there is a growing body of research which advocates treating phonological representation as being "substance free." Blaho (2008) provides a relatively comprehensive overview of theories which take a substance free approach. Broadly speaking, the substance free approach which is most compatible with the theory of phonetic change which I am advocating is one where there is no fixed or typical phonetic implementation for a phonological feature cross-linguistically. It would be impossible for me to accept the assumption that there is a *fixed* phonetic implemen-

tation of phonological features because I am investigating exactly those cases where the phonetic implementation *changes*. The assumption that there is a *typical* phonetic implementation simply appears to be unlikely, given that it would imply that there is a typical vowel system, deviations from which cost some kind of energy. There are some reasonable explanations which don't resort to phonological explanation for the sorts of phonetic distributions which are more common than others (Liljencrants and Lindblom, 1972; de Boer, 2001; Boersma and Hamann, 2008), and explaining the same phenomenon twice is unnecessary.

But moreover, if there were typical phonetic implementations for phonological representations, sound changes would be more complex to explain. Weinreich et al. (1968) outlined a number of problems to be solved in the study of sound change which still remain the core focus of sociolinguistics today. There is, for example, the Actuation Problem, which is a puzzle about how historical, social, and linguistic events converged such that a sound change was triggered in a particular dialect at a particular time, and not in all dialects, and not in this dialect at an earlier, or later time. Another is the Transition, or Incrementation Problem, which is a puzzle about how a sound change progresses continuously in the same direction over multiple generations. If there were typical phonetics for phonological representations, this would introduce an additional problem, which we could call the Maintenance Problem, which would be a puzzle about how once a sound change has become sufficiently advanced, why it doesn't revert back to the typical, or lower energy phonetic distribution. Taking the rotation of the short vowels in the Northern Cities Chain Shift (Labov et al., 2006) as an example, Labov (2010) argues that its actuation can be explained in terms of the historical event of the Erie Canal opening, and the linguistic context of the mixture of New York City /æ/ tensing and New England /æ/ tensing. Based on other studies (e.g. Tagliamonte and D'Arcy, 2009), it is most likely that the incrementation of the NCCS most likely occurred during the adolescence of speakers' lives. The Maintenance Problem would pose the question of why the NCCS has not gradually reverted back to the typical phonetics we would expect for the phonological features in the dialect, because there would presumably be a constant bias towards such a reversion either in acquisition or in speech production or perception, otherwise the notion of typical phonetics would be totally vacuous. Now, perhaps future research into

the phonology-phonetics interface will find that there *are* typical phonetics for a fixed set of universal phonological features, meaning Maintenance Problem has been a heretofore unexamined problem in sound change. For the time being, though, I will conclude that there are not typical phonetics for phonological features due to the fact that there are unidirectional sound changes.

Following the assumptions that there are no fixed or typical phonetics for phonological features to their logical conclusion would suggest that there is not a fixed or universal set of phonological features (Odden, 2006; Blaho, 2008; Mielke, 2008). Phonological features devoid of any phonetic information would be strictly formal and relational, and the nature of their relationship to phonetics would be, as Pierrehumbert (1990) said, semantic. The phonology-phonetics interface, then, would then relate formal phonological representation to its phonetic denotation. However, a fully articulated theory of radically substance free phonological features lies just outside what can be adequately argued on the basis of the data available to me, and is also not entirely necessary to achieve interesting results.

I'll be discussing the output of the phonology-phonetics interface mostly in terms of targets in F1×F2 space, largely because the data I'm working with are vowel formant measurements, not because this is a substantive claim about the nature of phonetic representation. The phonetic representation may actually be gestural targets and relative timing information, similar to the proposal of articulatory phonology (Browman and Goldstein, 1986), or perhaps even another alternative perceptual mapping, but since I don't have articulatory or perceptual data to bring to bear on the question, I will implicitly stick to F1×F2. When it comes to formalizing the relationship between a phonological feature and its phonetic realization, however, I'll refer instead to the phonetic dimension at issue. For example, back vowel fronting will play a major role in the dissertation, so when it is necessary to get more explicit about the implementation rules involved, I'll describe them as mapping to a target along the "backness" dimension, for which I'll be using F2 (and in some cases F2-F1) as a proxy for quantitative investigation. In addition, I'll be describing implementation in terms of implementation rules that have a phonological input and phonetic output, like (2.16) for example.

(2.16) [+low] $\rightsquigarrow$ 0.1 height

I'm using "⤳" in the implementation rule in part to emphasize that this is a qualitatively different sort of process than similarly formulated phonological rules, and in part to represent the imprecision in this formulation. I am only describing the interface in terms of rules for notational and expository convenience. In reality, the interface probably involves a complex system of non-linear dynamics, like those described by Gafos and Benus (2006). Importantly, however, I will be treating these phonetic implementation rules as being strictly translational, meaning their output is *insensitive* to local phonological context. This is largely to avoid making phonetic implementation rules too powerful, and the resulting theoretical framework too weak. For example, take the common phenomenon of pre-voiceless vowel shortening. If phonetic implementation could be sensitive to phonological context, there would be fully three different ways to account for pre-voiceless shortening: (i) a phonological process adds/changes a [long] or [short] feature on vowels when preceding voiceless consonants, (ii) a phonetic implementation rule sensitive to the following phonological voicing gives vowels a shorter phonetic target, (iii) phonetic gestural planning reduces the duration of vowels when preceding phonetically voiceless consonants. Since a combination of (i) and (iii) are already largely sufficient to account for phenomena like pre-voiceless vowel shortening, and already contentiously ambiguous, it doesn't seem necessary to further expand the power of phonetic implementation to also account for patterns like these.

Fully resolving *what* the phonetic representation is, and *how* it is derived from the phonological representation is beyond the scope of this dissertation, and also unnecessary in order to say at least some things about the relationship between phonetic change and phonological representation with certainty. To recap, the assumptions I'm making about the phonology-phonetics interface are:

(2.17) Phonological and phonetic representations are qualitatively different.

(2.18) The interface operates over the surface phonological representation.

(2.19) The interface operates over phonological features.

These assumptions are relatively simple, but still more explicit than a lot of research on phonetic change. They also lead to a number of important consequences, such as the fact that segments

which share phonological features should also share targets of phonetic implementation. Additionally, it should be the case that those properties of phonological representations which the interface can utilize for phonetic implementation should also be the observed units of phonetic change. For example, if the interface can *only* operate over individual phonological features, then we should expect phonetic change to *always* effect entire phonological natural classes. As a strong assumption, this would be incredibly useful in determining what the appropriate feature system of language ought to be. But I believe this strong assumption will be impossible to adhere to for all cases, meaning that the interface must also operate bundles of features, or perhaps over gestalt representations of segments as a whole. This will be discussed a bit further in Chapter 5.

Phonetic Alignment and Interpolation

I have separated the assignment of phonetic targets from the alignment and interpolation of those targets in my model of the grammar because they are conceptually distinct, although they may be implemented in one large step in reality, as suggested by Gafos and Benus (2006). At this step in the process, phonetic targets may experience some temporal displacement due to the phonetic alignment constraints in the language (Zsiga, 2000), and segments which are unspecified for certain targets may have gestures interpolated through them (Cohn, 1993). It is phonetic coarticulation at this level of representation that produces what I may occasionally call "phonetic effects." For example, in Chapter 4, there is extensive discussion of the effect of /l/ on the fronting of /uw/ and /ow/. If the measurable effect of /l/ on /ow/ is due to articulatory phasing relationships between the velar articulation of /l/ and the vocalic gesture of /ow/, then I would describe this as phonetic coarticulation, or a phonetic effect. On the other hand, if /l/ triggers featural changes on /ow/ in the phonology, producing different surface phonological representation which thus has a different phonetic target, then I would describe this effect as "phonological." Of course, distinguishing between these two radically different sources of differentiation is non-trivial, and is, in fact the topic of almost the entirety of Chapter 4.

Universal Phonetics

What I call "Universal Phonetics" are those properties of the speech signal which are well and truly non-cognitive, thus outside the domain of controllable variation. This will include both physiological and acoustic properties outside of speakers' control. For example, most (but not all, (Simpson, 2009; Zimman, 2013)) differences in average F1 and F2 between speakers, specifically men and women, can be attributed to differences in vocal tract length (see Figure 2.4). That proportion of the difference between men and women which is attributable to this physiological difference has everything to do with the physical properties of acoustics, rather than the cognitive properties of speakers' minds. Since this is a dissertation about language change, I will be focusing on the latter, because presumably neither the physics behind acoustics nor human anatomy has changed over the time course under examination here.
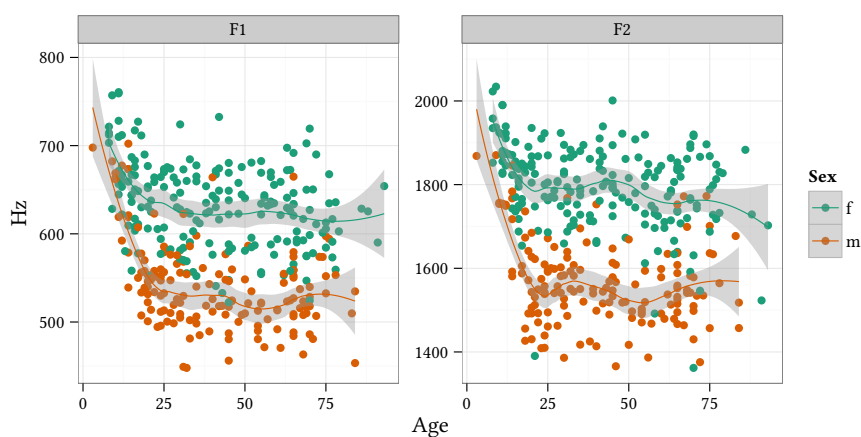
Figure 2.4: Mean F1 and F2 values by sex and age in unnormalized Hz.

2.2.3 Sociolinguistic Variation

In the architecture laid out above in §2.2.2, the role of sociolinguistic variation is not mentioned. I am following Preston (2004) in placing the "sociocultural selection device" outside of the core grammatical architecture. Rather, Preston (2004) and I posit that knowledge of sociolinguistic variation constitutes a separate and highly articulated domain of knowledge that utilizes optionality in the grammatical system. The way that utilization operates will, of course, depend on the

properties of the level of architecture under question. For example, choosing different phonological inputs, or phonological processes, will necessarily involve manipulation of the discrete and probabilistic properties of those systems, while altering the target of phonetic implementation will involve manipulation of the continuous properties of that system.

Constraining the range of options available to the sociocultural selection device to strictly those provided by the grammatical system is an important and principled move to make. For example, to my knowledge, it has never been reported for any speech community that speakers produce wh-island violations for sociostylistic purposes, and given the result from theoretical syntax that wh-island violations are a grammatical impossibility, we can go ahead and claim that they are also a sociolinguistic impossibility.

The scope of sociocultural selection device may also be broader than would be expected if it were an additional module of the grammatical system. By the modular feed-forward hypothesis, each module of the grammatical system can only make use of information passed to it by the preceding module. For example, when transforming phonological representations into targets for phonetic implementation, it should be the case that the interface can *only* be able to utilize surface phonological representations, and not, say, morphological information. However, MacKenzie and Tamminga (2012) have shown that patterns of variation are affected by factors which cannot trigger categorical grammatical processes. For example, the probability that an auxiliary will contract onto an NP subject is influenced by the length of the NP, but NP word length is not known to be a triggering factor in any categorical grammatical process. Tamminga (2012) has also demonstrated with a number of variable processes that choosing one grammatical option will boost the probability of choosing that same option again, with a decaying strength as the time lag between instances increases. Again, no categorical grammatical process appears to be triggered based on a combination of what happened at the last instance it could have applied, and how long ago that instance was. In the context of the grammatical architecture I've laid out here, it may be possible that the sociocultural selection device can look ahead and choose a phonological input on the basis of how it will be phonetically implemented, something that the grammatical system itself cannot do.

29

2.3   Phonetic Change

The focus of this dissertation will be on sound changes like those discussed in §2.1.4, which I will argue should be described as shifts in the phonetic implementation of surface phonological representations, as discussed in §2.2.2. In this section, I will discuss these kinds of changes in greater detail.

2.3.1   Phonetic Change is Continuous

For the purpose of this dissertation, I will use the term "phonetic change" to refer specifically to changes which progress continuously, in any fashion, through the phonetic space. There are some changes which may be called "phonetic" based on other principles, but they will not fall under this definition here. A good example is the shift in Montreal French from an anterior (apical) to a posterior (dorsal) version of /r/. Sankoff and Blondeau (2007) describe this change as progressing discretely, both in terms of the phonetics (tokens of /r/ were realized either as [r] or as [ʀ]) and in its progression through the speech community (most speakers used only one or the other variant). They also describe this as a phonetic change in /r/, because "the change in the phonetics of /r/ does not appear to interact with other aspects of Montreal French phonology" and "does not have systemic phonological consequences." This is a reasonable way to define "phonetic." This change in /r/ did not:

(2.20) alter the system of phonological contrasts, either by merging with an existing phoneme, or splitting to form a new one.

(2.21) alter the phonological grammar, either by ceasing or starting to be a target or trigger for any processes.

By these definitions, it was not a phonological change. However, it does not meet the definition of "phonetic change" that I will be using here, because of the categorical nature of the change. Presumably this change in /r/ *did* involve a shift in its natural class membership, joining the set of apical consonants.

On the other hand, there may be some discontinuities in sound changes that I would call phonetic. There is not empirical example of this, to my knowledge, but it is predicted to be possible under Quantal Theory (Stevens, 1989; Stevens and Hanson, 2010), whereby continuous shifts in articulation are related nonlinearly to acoustic realizations. That is, there are some regions in articulatory space where large differences correspond to relatively small acoustic differences, and other regions where small differences correspond to relatively large acoustic differences. A good example is the difference between bunched and retroflex articulations of /r/ in English. The two articulatory strategies for producing /r/ are drastically different, but correspond to only a very small acoustic difference in the distance between F3 and F4 (Espy-Wilson and Boyce, 1994).

Figure 2.5 displays a schematic diagram of the relationship between a hypothetical articulatory dimension and its corresponding acoustic realization. If there were a phonetic change progressing at a steady rate along the articulatory dimension, we would expect to observe a very slow rate of change in the acoustics (the measurable aspect of change for most studies) through the regions shaded in grey, with a sudden spike, or jump through the region in white.
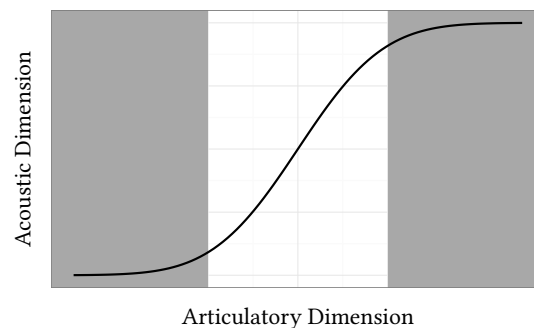

Figure 2.5: The proposed quantal relationship between changes in articulation and changes in acoustic realizations

Sharp discontinuities in the time course of any langauge change may also occur due to sociolinguistic reasons. For example, during the rise of *do*-support in early modern English, there was a brief period of time where the frequency of use of *do*-support dove sharply in the context of negation. Warner (2005) attributes this sharp effect to the development of a negative evaluation, and thus avoidance, of the form *don't*. This sociolinguistic influence generated a large perturba-

tion in the observed trajectory of the change, but does not force us to reevaluate the underlying grammatical analysis proposed for how this change progressed.

Simulating Phonetic Change as Categorical

However, it is still worthwhile to figure out whether phonetic changes which have appeared to progress continuously through the phonetic space could be simplified as the competition between two discrete phonetic targets. If this simplification could be done, it would have a number of desirable consequences. First and foremost, it would reduce the necessary complexity of the phonology-phonetics interface. The change of pre-voiceless /ay/ in Philadelphia from [ɑɪ] to [ʌi] could be described in terms of competing phonological representations without invoking language specific phonetic targets for their implementation. In fact, some frameworks of the phonology-phonetics interface do not allow for language specific phonetics, most notably Hale and Reiss (2008), where categorical phonological representations are "transduced" directly into articulatory gestures by the interface between the linguistic system and the biophysical system. "Transduction," according to Hale and Reiss (2008), involves no learning, and is part of humans' universal genetic endowment. This is, admittedly, the more parsimonious hypothesis on a number of conceptual grounds, and if it were also supported by the necessary empirical evidence it should be adopted. Second, the dynamics of phonetic change could be reduced to essentially the same ones that govern phonological, morphological, and syntactic change. The properties of competing discrete forms are fairly well understood within variationist work, and could be immediately imported for the purpose of understanding phonetic change.

As it stands, there has not been a rigorous attempt on the part of those studying phonetic change to demonstrate that it doesn't progress as competition between more-or-less categorical variants. For the most part, sociophonetic methodology involves the examination and statistical analysis of *means*. However, if this change were progressing as the categorical competition between two variants, merely examining the means would would not reveal this fact, and would actually make the change appear indistinguishable from continuous movement of a phonetic target through phonetic space. This fact is more obvious when looking at changes that must progress

in terms of categorical competition, like syntactic change. Figure 2.6 displays the loss of V-to-T movement in negative declarative sentences in Early Modern English as collected by Ellegård (1953). Each individual clause can only either have *do*-support, or have verb raising, as it would be impossible to, say, raise the verb 55% of the way to tense. Each point in the plot represents the proportion of *do*-support in an Early Modern English document.
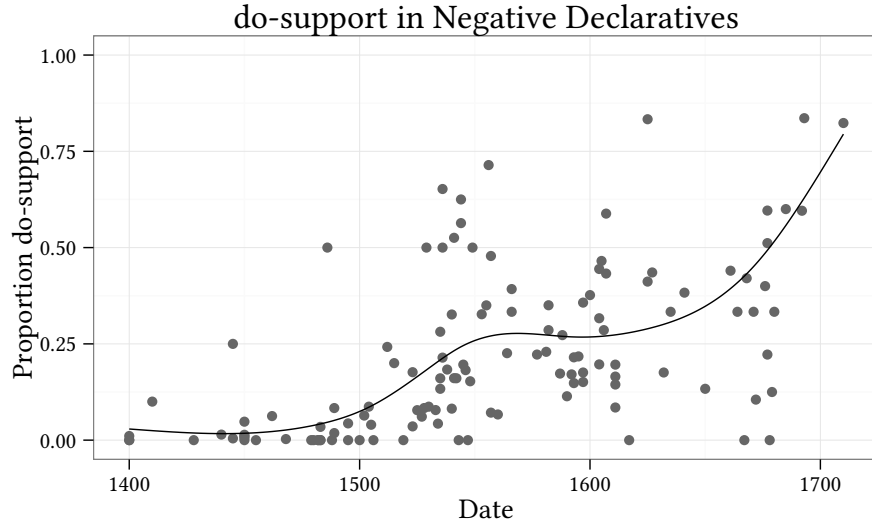

Figure 2.6: The loss of 'V-to-T' movement in Early Modern English

When coding tokens of *do*-support as 1 and tokens of verb raising as 0, the proportion of do-support for a given document is simply the mean of this sequence of 1's and 0's. A misinterpretation of Figure 2.6 would be that there was a continuous shift in *do*-support. Even though the average proportion of *do*-support changed gradually over time, any given token of *do*-support from any time point will still either be categorically verb raising, or tense lowering. It does not follow, then, that the diachronic trajectory of *means* reflects the synchronic pattern of variation. Looking at Figure 2.7, which depicts the raising of /ay/ in pre-voiceless contexts in Philadelphia, we cannot assume then that just because there is a continuous change in means along the diachronic dimension that the synchronic variation at any time point was also continuous.

However, there are other properties of the *distributions* of observations within speakers that can cast some light on whether or not this change progressed as categorical competition between
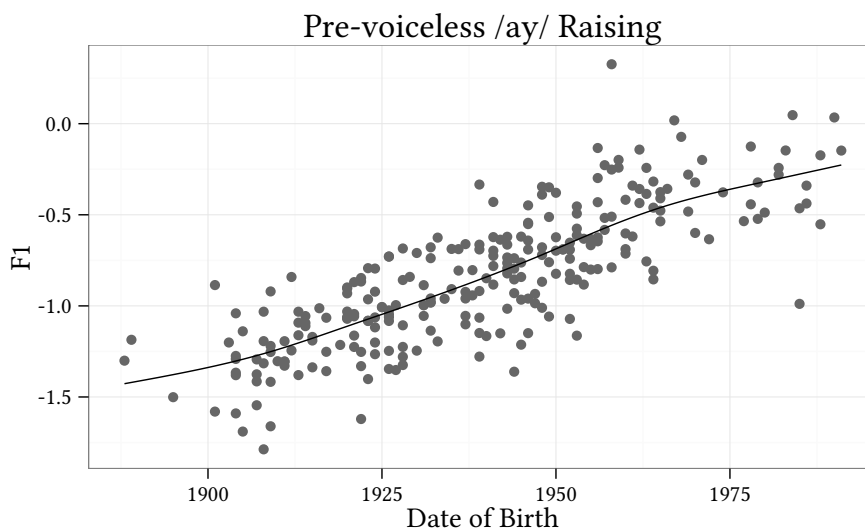
Figure 2.7: Pre-voiceless /ay/ raising.

[ɑɪ] and [ʌi], or whether it progressed in a phonetically gradual way between these two targets. The question essentially comes down to whether or not speakers' data is bimodal. Assessing whether or not data is multimodal, especially when we do not have any a priori basis for placing observations into categories, is statistically non-trivial. Some methods exist which rely mostly upon comparing the goodness of fit of a model which treats the data as monomodal, to a model which treats the data as bimodal. However, if the "truth" is that there are two modes, but their centers are close, and their variance broad, these tests will most likely fail to detect that fact. Moreover, these tests usually require more data than we have available per-speaker for sufficient power. Instead, here I will compare the observed data to the expected patterns from simulation in broad qualitative terms. The qualitative results are so striking and overwhelming that if there were a statistical null hypothesis test associated with them, statistical significance would be virtually guaranteed.

The distributional properties of each speaker that I will be examining are their *standard deviation* and the *kurtosis*. Roughly speaking, the standard deviation of a statistical distribution describes how broad the distribution is relative to its center. Kurtosis, on the other hand, describes

how peaked the distribution is. Figure 2.8 illustrates three different distributions which differ in their standard distributions and kurtosis. As a distribution becomes more broad, its standard deviation increases, and as it becomes more plateau-like, its kurtosis decreases. Darlington (1970) argued that kurtosis is actually best understood as a measure of bimodality, with low kurtosis indicating high bimodality, which makes it a perfect measure for the problem at hand.
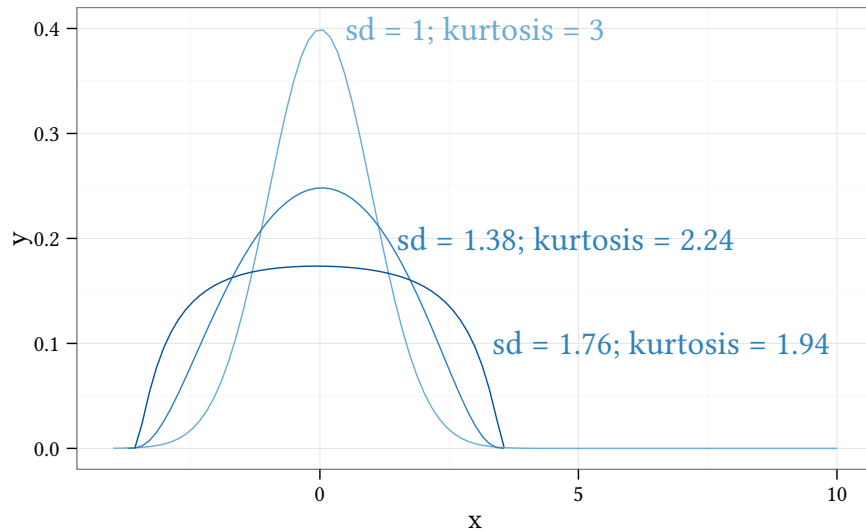

Figure 2.8: Three distributions differing in standard deviation and kurtosis.

When mixing two distributions, there will be a systematic relationship between the mean of the mixture and its standard deviation and kurtosis. Figure 2.9 illustrates what phonetic change which progresses as competition between two categorical variants would look like. Each facet represents one hypothetical speaker who varies in choosing category A or B with some probability. The label for each facet represents the probability that the speaker will choose variant A. Category A has a mean of 1.5 and a standard deviation of 1, while Category B has a mean of -1.5, and a standard deviation of 1. The phonetic targets for Categories A and B are the same for all speakers; all that differs between speakers is the mixture proportions of A and B.

While the fundamental behavior of these speakers is categorical and probabilistic, given the relative closeness of the phonetic targets for Categories A and B, a researcher would not be able to tell on a token by token basis which category a speaker intended to use in a particular instance.
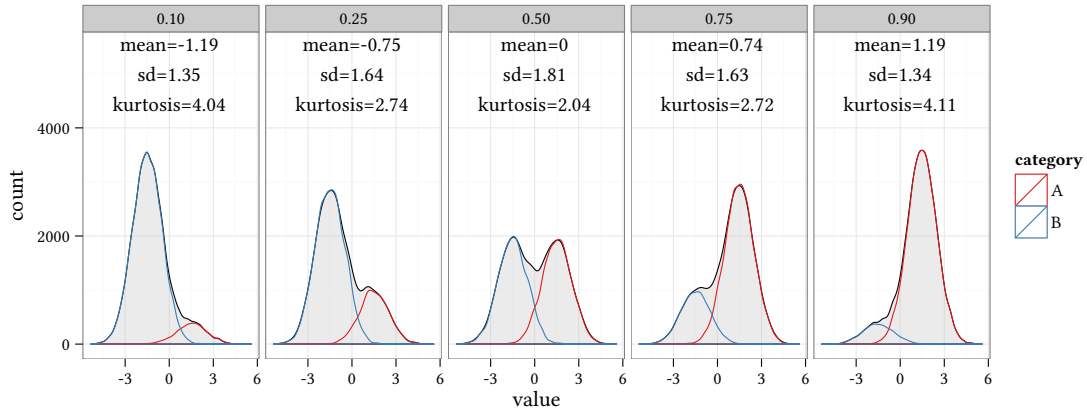
Figure 2.9: An illustration of the systematic relationship between mean, standard deviation, and kurtosis of the mixture of two distributions.

Thus, all that is observable to the linguist is the over-all distribution of the mixture of the two categories, represented by the shaded regions in Figure 2.9. However, as is annotated in each facet of Figure 2.9, there is a systematic relationship between the mean of the mixture distribution and its standard deviation and kurtosis. The more homogeneous mixtures (the far left at 0.1 and far right at 0.9) have the most extreme means, fairly close to just the pure means of just Category A and Category B. They also have the lowest standard deviations and highest kurtosis. The most even mixture (the center, at 0.5) has a mean that is almost exactly in the middle between Category A and B. It also has the broadest distribution, giving it the highest standard deviation, and is the most plateau-like, giving it the lowest kurtosis.

If phonetic change progressed as competition between categorical variants, then we should expect to see a systematic relationship between the mean of speakers' data and the standard distribution and kurtosis of their data. The raising of pre-voiceless /ay/ in Philadelphia is perhaps the perfect example of phonetic change to examine for this kind of relationship. First, the change progressed mostly along just one dimension: F1. Second, it covered a very large range of F1 values from beginning to end, starting off with an essentially low nucleus and ending with an essentially mid one. We can make a principled argument that there is a phonological difference between these two endpoints ([+low] to start, [−low] to end), and the phonetic difference between them

36

is large enough that we ought to observe as strong a relationship between the mean, standard deviation and kurtosis as we could expect to for any phonetic change.

Figure 2.10 plots speakers' mean F1 values against the standard deviation and kurtosis of F1. As a first pass attempt to look for a systematic relationship between means, standard deviation and kurtosis, this figure does allow much hope of finding one. The standard deviation of speakers' F1 is strikingly consistent across the entire range of F1 means, and the mixture hypothesis would predict a marked peak in the middle. The kurtosis of speakers' F1 values is also very flat, and on average slightly larger than 3, which is the kurtosis for a normal distribution. The mixture hypothesis would predict a marked drop in kurtosis in the middle of the F1 range.
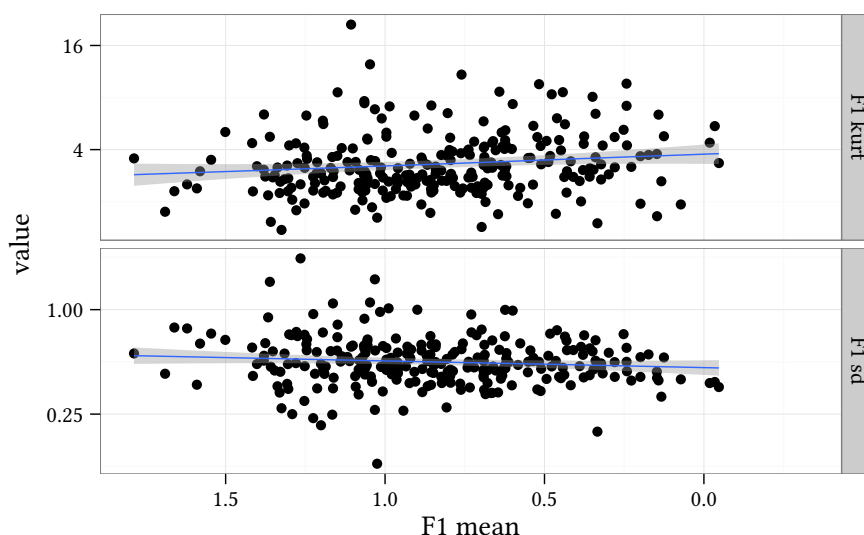

Figure 2.10: Comparing speakers' means to their standard deviation and kurtosis for /ay0/.

It is possible to generate more precise expectations about what the standard deviation and kurtosis of mixtures of [ɑɪ] and [ʌi] would be through simulation. Figure 2.11 displays the distribution of data for the 4 most conservative and 4 most advanced [ay0] speakers in the corpus. Briefly assuming that the /ay/ raising progressed as categorical variation between [ɑɪ] and [ʌi], we can also assume that these extreme speakers have relatively pure mixtures of just one or the other variant simply because their data lie on the extremes. We can sample tokens from these two sets of speakers at different mixture rates to simulate new speakers that lie along the continuum

from conservative to innovative. The distributional properties of these simulated speakers should roughly approximate the expected distributions of speakers for whom /ay/ raising progresses as categorical competition.
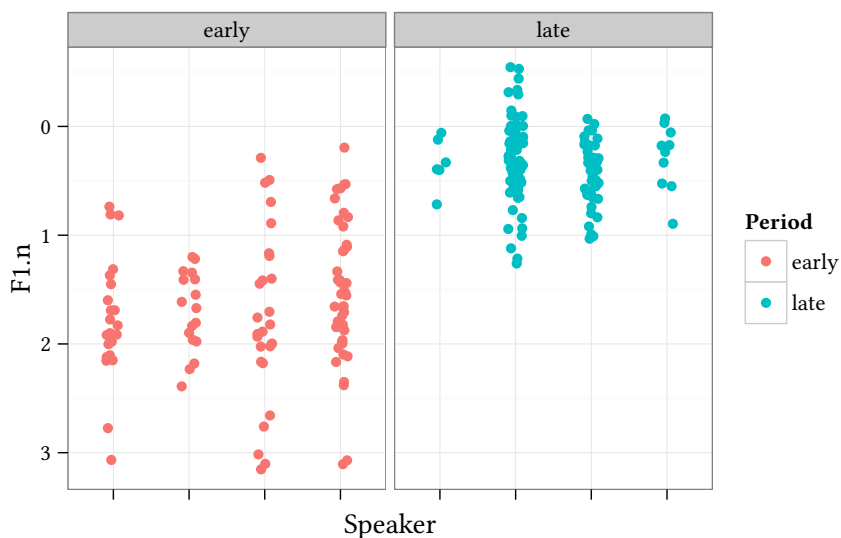

Figure 2.11: Distribution of [ay0] data for the 4 most conservative and 4 most advanced speakers.

For these simulations, I capped the maximum number of tokens that a single real speaker could contribute to the pool of tokens I'd resample from at 30. For every simulated speaker, I sampled 40 tokens from the original speakers' data with replacement. The proportion of tokens sampled from the conservative ([ɑɪ]) vs innovative ([ʌi]) pool varied from 0%:100% all the way to 100%:0% by increments of 1%. For each mixture proportion, I simulated 100 speakers. Figure 2.12 plots the data from 9 simulated speakers at different mixture rates. The far left facet displays simulated speakers who drew from the innovative pool of data 10% of the time, the middle facet displays simulated speakers who drew from the innovative pool 50% of the time, and the far right facet simulated speakers who drew from the innovative pool 90% of the time.

Figure 2.13 plots the relationship between mixture proportions of these 9 simulated speakers and their distributional properties, specifically F1 mean, standard deviation, and kurtosis. As was necessarily going to be the case, as the mixture of of innovative variants increases, the mean F1 drops (raising /ay/ in the vowel space). The most even mixtures of conservative and innovative
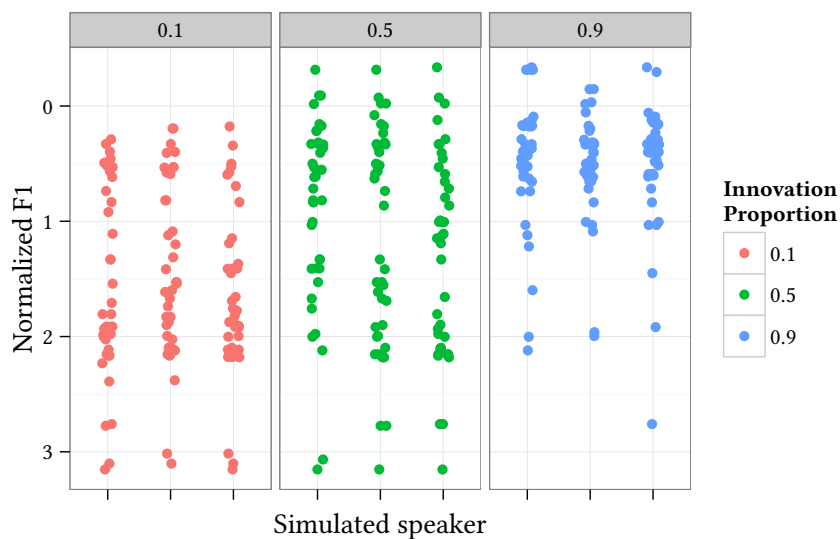
Figure 2.12: Simulated speakers at the beginning, middle, and end of the change.

variants have the largest standard deviation, and the smallest kurtosis.
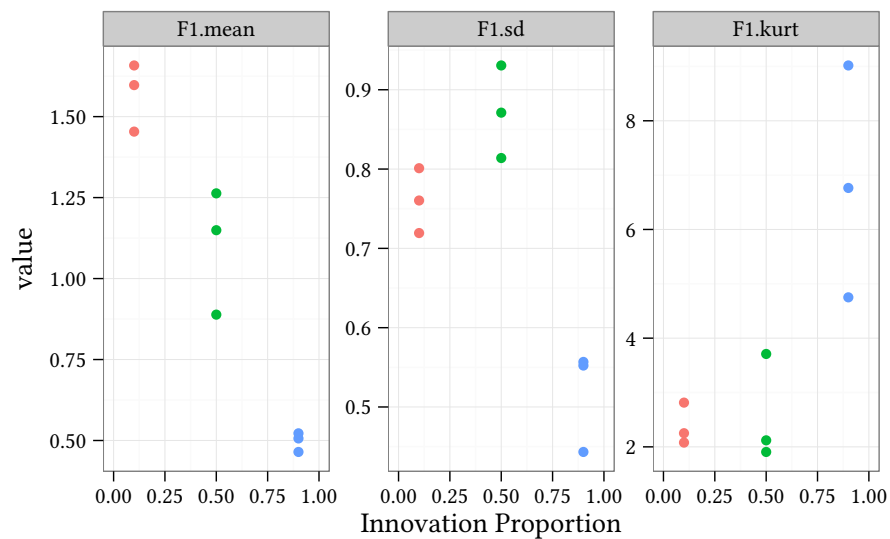

Figure 2.13: The effect of mixing distributions on three different diagnostics: mean, standard deviation, and kurtosis.

The mixture proportion of conservative and innovative variants of real PNC speakers is un-

known (and as we'll see, is not actually how this change is progressing). However, since the relationship between mixture proportion and mean F1 is linear and monotonic (as seen in the left facet of Figure 2.13), we'll compare the mean F1 to the other distributional properties of real speakers and of the simulated speakers.

Figure 2.14 displays the first of these comparisons, plotting the mean of F1 against the kurtosis of F1. The filled blue contours represent the region of highest density for the simulated speakers, and the blue line is a cubic regression spline fit to the simulated data. As expected, the simulated speakers have a dip in the kurtosis, indicating more bimodality, about midway through the course of the change. The red points represent the data from real PNC speakers, and the red line is a cubic regression spline fit to their data. Many real speakers fall within the high density regions of the simulated speakers, but the over-all relationship between mean F1 and F1 kurtosis is totally different. While the simulated speakers have a kurtosis well below that of a normal distribution (represented by the horizontal black line) midway through the change at F1 means slightly less than 1, the real speakers' kurtosis is, on average, slightly larger than a normal distribution. This means that simulated speakers have very plateau-like distributions to their data midway through the change, while real speakers actually have rather peaked distributions throughout the change, including the midpoint.

Figure 2.15 plots the second key relationship between mean F1 and F1 standard deviation. Again, the blue contours represent the region of highest density for simulated speakers, and the blue line is a cubic regression spline fit to the simulated speakers. Again, the red points represent the data of real speakers from the PNC, and the red line a cubic regression spline fit to their data. The mismatch between simulated expectations and real data is even more striking in this case. Almost no real speakers have the standard deviation of F1 we would expect at almost every stage of the change. In fact, the standard deviation of F1 across speakers remains remarkably stable throughout the change.

The conclusion we can draw is that the model of phonetic change whereby /ay/ raised from [ɑɪ] to [ʌi] through categorical variation between these two forms is a poorly fitting one. Rather the fact that both the standard deviation of F1 and its kurtosis remains essentially constant
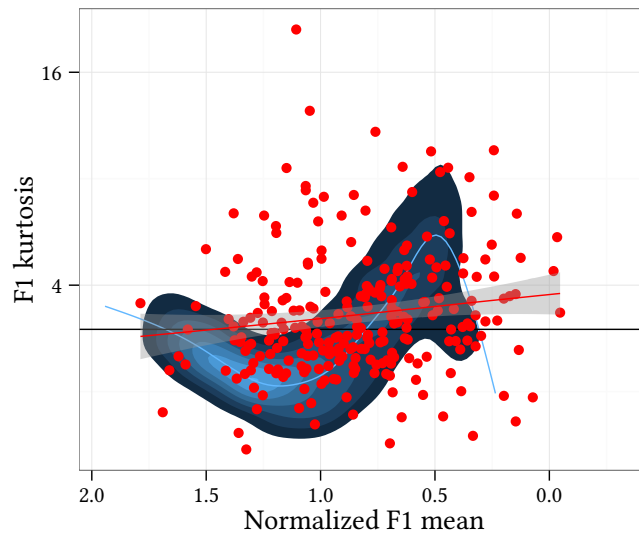
Figure 2.14: The relationship between normalized F1 mean and kurtosis as observed in speakers, overlaid on the two dimensional density distribution from the mixture simulation. Note the y-axis is logarithmic. The horizontal line at kurtosis=3 represents the kurtosis of a normal distribution.
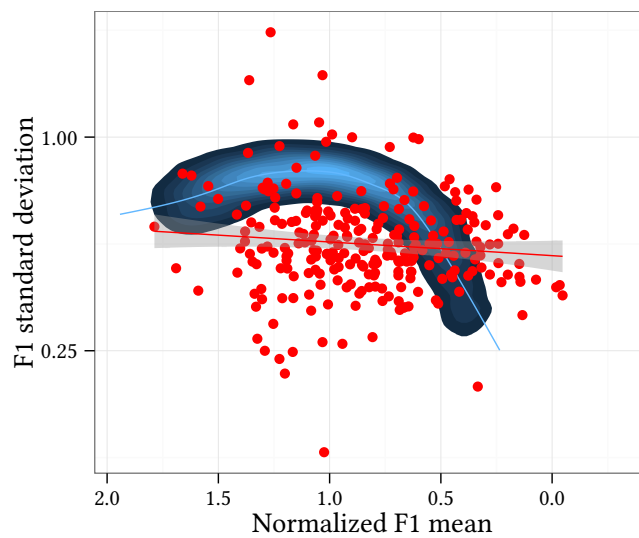

Figure 2.15: The relationship between normalized F1 mean and standard deviation as observed in speakers, overlaid on the two dimensional density distribution from the mixture simulation. Note the y-axis is logarithmic.

throughout the change, with only the mean changing, lends support to the model where /ay/ raised to a mid position through gradual phonetic change of a single phonetic target. This model of phonetic change, which has actually been the default assumption of sociolinguists for good reason, necessitates language specific phonetic implementation, for the reasons laid out in the beginning of this chapter. Language change is necessarily a change in speakers' knowledge of their language. This change progressed as continuous movement of a single allophone through the phonetic space, meaning speakers must have some kind of non-trivial phonetic knowledge which they acquired with the rest of their linguistic knowledge, and represented in some way. Based on the phonetics/phonology architecture laid out in §2.2.2, the most plausible locus of this knowledge is in the rules of phonetic implementation of phonological representations.

2.4   Conclusion

In this chapter, I have attempted to outline what is at stake, in terms of the architectural theory of phonetics and phonology, when diachronic analysis is brought to bear on the problem. The basic goal of modern linguistics is to understand what constraints there are on possible languages. Given that during language change from state A to state B, every intermediate state is *also* a language, then it follows that the path of language change is also constrained at all points by the same constraints as synchronic languages. So careful analysis of how language changes can inform our theories of synchronic grammar, and vice versa.

I have also tried to carefully define the particular object of study in this dissertation. "Phonetic change" is a phenomenon, but as I believe was made clear in §2.1.6, the outcomes of language change, like "merger" or "metathesis," are not unitary phenomena, but can arise through multiple different kinds of change to speakers' competence. The remainder of this dissertation will be devoted to supporting the primary claim of §2.3 that most of the observed phenomena related to "phonetic change" can be attributed to changing knowledge of the phonetic implementation of phonological representations, but also to determining which properties should be attributed to other domains of knowledge.

The results from §2.3.1 may be seen as suggestive that a categorical phonological represen-

tation is not necessary to capture the observed properties of phonetic change. However, this is not my conclusion, and the following chapters will also be devoted to demonstrating that *both* phonological and phonetic representations are necessary to capture the facts of sound change.

Bibliography

Anttila, Arto, Vivienne Fong, Štefan Beňuš, and Jennifer Nycz. 2008. Variation and opacity in Singapore English consonant clusters. *Phonology* 25:181. URL `http://www.journals.cambridge.org/abstract_S0952675708001462`.

Bailey, Charles-James. 1973. Variation and Linguistic Theory. Technical report, Center for Applied Linguistics, Washington.

Becker, Kara. 2010. Regional Dialect Features on the Lower East Side of New York City: Sociophonetics , Ethnicity , and Identity. Ph.d., New York University.

Becker, Kara, and Amy Wing-mei Wong. 2010. The Short-a System of New York City English: An Update. *University of Pennsylvania Working Papers in Linguistics* 15. URL `http://repository.upenn.edu/pwpl/vol15/iss2/3/`.

Bermúdez-Otero, R. 2007. Diachronic Phonology. In *The Cambridge Handbook of Phonology*, ed. Paul de Lacy, chapter 21, 497–517. Cambridge: Cambridge University Press.

Blaho, Sylvia. 2008. The Syntax of Phonology: A radically substance free approach. Ph.d., University of Tromsø.

Blevins, Juliette, and Andrew Garrett. 2004. The Evolution of Metathesis. In *Phonetically Based Phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade, chapter 5, 117—-156. New York: Cambridge University Press.

de Boer, Bart. 2001. *The Origins of Vowel Systems*. New York: Oxford University Press.

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

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

Buckley, Eugene. 2011. Metathesis. In *The Blackwell Companion to Phonology, Volume 3*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, volume 1968. Blackwell.

Bybee, Joan. 2007. Word Frequency in Lexical Diffusion and the Source of Morphophonological Change. In *Frequency of use and the organization of language*, ed. Joan Bybee, chapter 2, 23—-34. New York: Oxford University Press.

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

Coye, D. F. 2010. Dialect Boundaries in New Jersey. *American Speech* 84:414–452. URL `http://americanspeech.dukejournals.org/cgi/doi/10.1215/00031283-2009-032`.

Darlington, Richard B. 1970. Is Kurtosis Really "Peakedness?". *The American Statistician* 24:19–22.

Disner, Sandra Ferrari. 1978. Vowels in Germanic Languages. Ph.d., UCLA.

Dresher, B Elan. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.

Dresher, Elan, Christopher Harvey, and Will Oxford. 2012. Contrast Shift as a Type of Diachronic Change.

Durian, David. 2012. A New Perspective on Vowel Variation across the 19th and 20th Centuries in Columbus, OH. Ph.d., The Ohio State University.

Ellegård, Alvar. 1953. *The auxiliary do: The establishment and regulation of its use in English.*. Stockholm: Almqvist and Wiksell.

Espy-Wilson, Carol, and Suzanne Boyce. 1994. Acoustic differences between "bunched" and "retroflex" variants of American English /r/. *The Journal of the Acoustical Society of America* 95:2823.

Fourakis, Marios, and Robert Port. 1986. Stop Epenthesis in English. *Journal of Phonetics* 14:197–221.

Fruehwald, Josef, Jonathan Gress Wright, and Joel Wallenberg. forthcoming. Phonological Rule Change: The Constant Rate Effect. In *The proceedings of the North-Eastern Linguistic Society (NELS)*. URL `http://www.ling.upenn.edu/~joseff/papers/FGW_CRE_NELS40.pdf`.

Gafos, Adamantios I, and Stefan Benus. 2006. Dynamics of phonological cognition. *Cognitive Science* 30:905–943. URL `http://onlinelibrary.wiley.com/doi/10.1207/s15516709cog0000_80/abstract`.

Gress-Wright, Jonathan. 2010. Opacity and Transparency in Phonological Change. Ph.d., University of Pennsylvania.

Guy, Gregory R. 2007. Lexical Exceptions in Variable Phonology. *University of Pennsylvania Working Papers in Linguistics* 13:109–119.

Hale, Mark. 2004. Neogrammarian Sound Change. In *The Handbook of Historical Linguistics*, ed. Brian D Joseph and Richard D Janda. Blackwell.

Hale, Mark, Madelyn Kissock, and Charles Reiss. 2007. Microvariation, variation, and features of universal grammar. *Lingua* 117:645–665.

Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. New York: Oxford University Press.

Hall, Daniel Currie. 2007. The Role and Representation of Contrast in Phonological Theory. Ph.d., University of Toronto.

Herold, Ruth. 1990. Mechanisms of merger: The implementation and distribution of the low back merger in eastern Pennsylvania. Ph.d., University of Pennsylvania.

Hoenigswald, Henry. 1960. *Language change and linguistic reconstruction*. University of Chicago Press.

Hume, Elizabeth. 2000. Metathesis Website. URL `http://www.ling.ohio-state.edu/~ehume/metathesis/`.

Johnson, Daniel Ezra. 2007. Stability and change along a dialect boundary: The low vowels of Southeastern New England. Doctoral Dissertation, University of Pennsylvania.

Keating, Patricia A. 1985. Universal phonetics and the organisation of grammars. In *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, ed. Victoria A Fromkin, 115–132. New York: Academic Press.

Keating, Patricia A. 1988. Underspecification in phonetics. *Phonology* 275–292.

Keating, Patricia A. 1990. Phonetic representations in a generative grammar. *Journal of phonetics* .

Kiparsky, Paul. 1965. Phonological Change. Ph.d., Massachusetts Institute of Technology.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.

Kroch, Anthony. 1994. Morphosyntactic Variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K Beals.

Labov, William. 1969. Contraction, Deletion, and Inherent Variability of the English Copula. *Language* 45:715–762.

Labov, William. 2001. *Principles of linguistic change. Volume 2: Social Factors. Language in Society*. Oxford: Blackwell.

Labov, William. 2010. *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors*. Oxford: Blackwell.

Labov, William, Sherry Ash, and Charles Boberg. 2006. *The Atlas of North American English*. New York: Mouton de Gruyter.

Labov, William, and Maciej Baranowski. 2006. 50 Msec. *Language Variation and Change* 18:223–240.

Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear Incrementaion, Reversal, and Reanalysis. *Language* 89:30–65.

Liberman, Mark, and Janet B Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In *Language sound structure: studies in phonology presented to morris halle*, ed. Mark Aronof and R T Oehrle, 157–233. Cambridge: MIT Press.

Liljencrants, Johan, and Björn Lindblom. 1972. Numerical Simulation of Vowel Quality Systems : The Role of Perceptual Contrast. *Langauge* 48:839–862.

MacKenzie, Laurel, and Meredith Tamminga. 2012. Non-local Conditioning of Variation: Evidence and Implications.

Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. New York: Oxford University Press.

Mohanan, K. P. 1992. Describing the phonology of non-native varieties of a language. *World Englishes* 11:111–128. URL `http://doi.wiley.com/10.1111/j.1467-971X.1992.tb00056.x`.

Odden, David. 2006. Phonology ex nihilo. Paper presented at the the Tromsø Phonology Project Group Meeting.

Oxford, Will. 2012. Patterns of contrast in phonological change : Evidence from Algonquian vowel systems. URL `http://individual.utoronto.ca/woxford/Oxford_Ms_PatternsOfContrast.pdf`.

Parrell, Benjamin. 2012. The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of phonetics* 40:37–45. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3381366&tool=pmcentrez&rendertype=abstract`.

Phillips, B S. 1999. The mental lexicon: evidence from lexical diffusion. *Brain and language* 68:104–9. URL `http://www.ncbi.nlm.nih.gov/pubmed/10433746`.

Phillips, Betty S. 1984. Word Frequency and the Actuation of Sound Change Betty S. Phillips. *Language* 60:320–342.

Phillips, Betty S. 2006. *Word Frequency and Lexical Diffusion*. Basingstoke: Palgrave Macmillian.

Pierrehumbert, Janet B. 1990. Phonological and phonetic representation. *Journal of Phonetics* 18:375–394.

Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34:516–530. URL `http://linkinghub.elsevier.com/retrieve/pii/S009544700600043X`.

Preston, Dennis R. 2004. Three kinds of sociolinguistics: A psycholinguistic perspective. In *Sociolinguistic variation: Critical reflections*, ed. Carmen Fought, 140—-158. New York: Oxford University Press.

Ruch, Hanna. 2012. Investigating a gradual metathesis: Production and perception of /s/ aspiration in Andalusian Spanish.

Sankoff, Gillian, and Hélène Blondeau. 2007. Language Change across the Lifespan: /r/ in Montral French. *Language* 83:560–588.

Scobbie, James M. 2005. The phonetics phonology overlap. *QMU Speech Science Research Centre Working Papers, WP-1* 1–30.

Simpson, Adrian P. 2009. Phonetic differences between male and female speech. *Language and Linguistics Compass* 3:621–640. URL `http://doi.wiley.com/10.1111/j.1749-818X.2009.00125.x`.

Stevens, Kenneth. 1989. On the *quantal* nature of speech. *Journal of Phonetics* 17:3–45.

Stevens, Kenneth N, and HM Hanson. 2010. Articulatory-Acoustic Relations as the Basis of Distinctive Contrasts. In *The Handbook of Phonetic Sciences*, ed. William J. Hardcastle, John Laver, and Fiona E. Gibbon, 424–453. Blackwell, second edition.

Tagliamonte, Sali, and Alexandra D'Arcy. 2009. Peaks Beyond Phonology: Adolescence, Incrementation, and Language Change. *Language* 85:58–108.

Tamminga, Meredith. 2012. Persistence in the production of linguistic variation.

Torreira, Francisco. 2006. Coarticulation between Aspirated-s and Voiceless Stops in Spanish : An Interdialectal Comparison. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 1978, 113–120. Somerville, MA: Cascadilla Proceedings Project.

Torriera, Francisco. 2007. Pre- and postaspirated stops in Andalusian Spanish. In *Segmental and prosodic issues in Romance phonology*, ed. P. Prieto, J. Mascaró, and M.-J. Solé, 67–82. Philadelphia: John Benjamins.

Warner, Anthony. 2005. Why DO dove: Evidence for register variation in Early Modern English negatives. *Language Variation and Change* 17:257–280.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics*, ed. W Lehmann and Y Malkiel. U. of Texas Press.

Zimman, Lal. 2013. Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality* 2:1–39. URL `http://openurl.ingenta.com/content/xref?genre=article&issn=2211-3770&volume=2&issue=1&spage=1`.

Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28:69–102.