

Phonological Involvement in Phonetic Change

Josef Fruehwald

1 Introduction

A long standing question in sound change is what role phonology plays. For example, the neogrammarian controversy revolved around whether sound change progressed word-by-word (lexical diffusion), or at once across all potential contexts (neogrammarian sound change) (Labov, 1981). The latter implicates some process of generalization, which could be described as “phonological.”¹ The issues involved in the debate between generalized sound change and item-by-item sound change reach far beyond historical linguistics. For example, some early and influential proposals in exemplar theory suggested that stable generalizations of the sort necessary for neogrammarian sound change are absent from language; rather, they are created at speaking time, then lost again. In my dissertation, I will explore the generalized, phonological nature of many phonetic changes, and attempt to address objections from lexical diffusion and exemplar theory.

Another vexing aspect of neogrammarian sound change is its phonetic graduality. The very fact that many sound changes progress in a phonetically gradual way (rather than a phonetically abrupt, but variable way), is challenging for approaches to the phonology-phonetics interface which don’t allow for language specific phonetics. It also sets phonetic change apart from most other attested forms of language change,² which progress as discrete variation. However, even for those who do allow for language specific phonetics, it is unclear when phonological reanalysis takes place. For example, a vowel which had been historically designated phonologically [+hi] may begin to lower phonetically. Eventually, the [+hi] specification would become implausible, given the degree of phonetic lowering, and a [−hi] specification would become more likely.

Finally, there are the cases of conditioned sound changes which result in categorical phonological processes at their end points (at the latest). There has been a lot of

¹Other descriptions of neogrammarian generalization will be discussed later.

²Phonetic change may share gradualness with semantic change, cf. the quote from Weinreich et al. (1968) below.

speculation about how gradient “phonetic effects” of this sort transition into phonological processes. The consensus view holds that the phonetic effect of a conditioning environment gradually increases in its magnitude until such time that it is accidentally reanalyzed as being phonological in origin (e.g. hypo/hyper-correction (Ohala, 1981), Evolutionary Phonology (Blevins, 2004), early stages in the Lifecycle of Phonological Change (Bermúdez-Otero, 2007), and many approaches to Exemplar Theoretic sound change, like Bybee (2002)). However, most work supporting the consensus view is based on the outcome of historical changes, perceptual experiments, conceptual argumentation, and simulation, not on the direct study of sound change in progress. My focus on sound change in progress is another unique aspect of this dissertation proposal. I will certainly turn to conceptual argumentation and simulation, but these will always be checked against rich data on real language change.

All of these points of inquiry into phonetic change could be synthesized into the following question: *When* does phonology play a role in phonetic change? This is the question that I propose to pursue in my dissertation. My preliminary results suggest that the answer is “early, and continuing.”

Early...

By “early,” I mean that phonological generalizations and reanalyses occur either at the very onset of a phonetic change, or at least earlier than previously expected. In this proposal, I’ll support the early effects of phonology on phonetic change with two case studies of conditioned sound change in Philadelphia: /ay/ raising in pre-voiceless environments, and /ey/ peripheralization before consonants.

...and continuing.

By “continuing,” I mean that at every point during a phonetic change, it is mediated by phonological representation, and phonological contingencies. I support this assertion with two analyses. The first is an analysis of parallel vowel shifts, specifically recent descriptions of the Canadian Shift. The second is another analysis of /ay/ in Philadelphia. It appears that as /ay[–voice]/ is raising phonetically, /ay[+voice]/ is becoming shorter. That is, there appears to be some cue-trading with regards to the phonological contrast as a consequence vowel quality change in /ay[–voice]/.

Both of these aspects, “early” and “continuing,” are non-trivial and, I think, unique empirical and theoretical contributions of the proposed dissertation. The point at which I am proposing abstract phonological conditioning on phonetic change exists is far earlier than the consensus view. With “continuing,” I am proposing a direct and continuing phonological influence on phonetic change which I intend to make explicit.

1.1 The Philadelphia Neighborhood Corpus

To attempt to address these questions, I’ll draw heavily from the Philadelphia Neighborhood Corpus (hereafter PNC) (Labov and Rosenfelder, 2011). It is based on the catalogue of interviews conducted by members of the Ling 560 class at the University of Pennsylvania between 1973 and 2010. I myself was a member of the class in 2006. Utilizing a modified version of the Penn Phonetics Lab Forced Aligner (Yuan and Liberman, 2008), and a modified version of automated, Bayesian formant tracking which was originally developed by Keelan Evanini for his dissertation (Evanini, 2009), the PNC consists of nearly 800,000 measurements from 327 speakers.

The following features make the PNC a novel resource for the study of sound change.

Large

The PNC is relatively large compared to other Sociolinguistic corpora. The Buckeye Corpus (Pitt et al., 2007), for example, is another corpus of sociolinguistic interviews which is also phonetically aligned. However, it has only 40 speakers, compared to the PNC’s 327. There are other large sociolinguistic corpora of large urban centers, like the Montreal corpus (Sankoff and Sankoff, 1973) with 120 speakers from the original study, and the Toronto English Corpus with 214 speakers, but they have not been subjected to the same degree of phonetic analysis.

A large corpus like the PNC makes it possible to do more in-depth investigation of both linguistic and social conditioning of sound change that would be impossible with a smaller corpus. The principle that most linguistic contexts appear rarely (Zipf’s Law) means that data which is rare in the PNC would be completely absent in a smaller corpus.

Long

The PNC has an unparalleled time depth. Interviews were conducted between 1973 and 2010 (and will be conducted again in 2012), leading to about 40 years of real time coverage of the Philadelphia speech community. Assuming that speaker’s linguistic behavior

is largely representative of the speech community of their youth (an apparent time analysis), the time depth is now expanded to include over a century of data, with speakers born between 1888 and 1991. This broad time depth means that it is possible to observe sound changes throughout most of their lifespan, and fewer time points are up to guess work and speculation.

Focused

Among large speech corpora, the PNC is also unique in being carefully focused on just one speech community. It also contains more detailed demographic data than most large speech corpora, including age, ethnicity, years of education, and neighborhood.

Expandable

Perhaps most importantly, the PNC is *expandable*, due to the automated tools which were used to create it. Additional interviews that I conduct, and other speech samples that I find, or that are contributed by other researchers, are easily added to the corpus once they have been transcribed. Even though it is already possible to do so much with the corpus as it now stands, it is not static, or completed, in any sense.

2 Phonetic Change vs All Other Language Change

For being so well studied for so long, it is surprising that it is not commented on more often how unique phonetic change is compared all other kinds of language change and variation. In fact, one of Weinreich's general criticism of Herman Paul's work was that it focused too narrowly on phonetic change.

Had Paul been interested in the problem of discreteness versus continuity as a feature of language design, he might have both enlarged on the parallelisms between *phonetic and semantic change*, and realized that *they are jointly not representative of the rest of language change*. (Weinreich et al., 1968, italics added)

A whole host of syntactic, morphological, and phonological variables are describable in terms of variation between two competing outcomes. Variation of this sort was first formalized for phonology in terms of variable rules. The early formulation of variable rules in Labov et al. (1968) and Labov (1969) were redescribed by Cedergren and Sankoff (1974) using logistic regression, launching the long and fruitful VARBRUL paradigm.

More recently, Kroch (1989) formulated discrete variation between two alternatives in terms of grammar competition, where G_A produces variant A, and G_B produces variant

B. The grammar competition paradigm has proven to be very fruitful in the field of Historical Syntax. Yang (2000, 2002) combined principles drawn from biological dynamics with reinforcement learning of competing grammars to produce a model that accurately accounts for both patterns in language acquisition and historical change. To simplify Yang's model, if grammar G_A produces more strings than G_B can't parse than vice versa, G_A will eventually replace G_B .

Competing grammars analysis has even proven to be fruitful in the domain of sound systems. Fruehwald et al. (2009) established that the data on the loss of word-final devoicing in Early New High German was well fit by the competing grammars model. One grammar contained the word final devoicing rule, and the other did not. Fruehwald et al. (2009) diagnosed the validity of the competing grammars model by investigating the Constant Rate Effect. There were not significant differences in the rate of change between different segments within the dialects of Early New High German under investigation, indicating that the variation was between a process which targeted all segments, and no process at all. Figure 1 plots the data from 5 dialects which were studied.

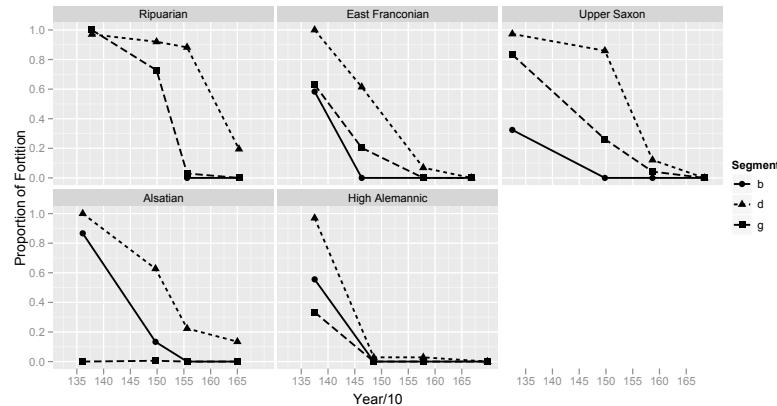


Figure 1: Constant rate effect in phonological change.

If it were possible to describe phonetic change in terms of competition, it would certainly be a "good thing" in a number of ways. First, it would mean phonetic change is simply another instance of a known type of language change: competitive change. We could then immediately bring the well established analytical and theoretical tools already developed for competitive change to bear on issues in phonetic change. Secondly, it would eliminate the necessity to appeal to language specific phonetic implementation. Competitive change could be handled sufficiently with phonological representation. In short, by describing phonetic change as competitive change, we would avoid "multiplying entities", and on conceptual grounds, one kind of change is better than two.

However, it does not appear that it is possible to avoid multiplying entities in this case. Figure 2 displays the rise of *do*-support in Middle English, a well known competitive change. Figure 3 displays the raising of /ay/ in pre-voiceless contexts in contemporary Philadelphia, a phonetic change identified as “new and vigorous” in the 1970’s (Labov, 2001). The apparent similarities between these two changes are strictly due to the graphical elements used to display them. Figure 2 displays the change in the proportion of *do*-support over time. That is, the normalized *frequency* of use of *do*-support. Figure 3 on the other hand, does not display the frequency of use of a raised variant, but rather the shifting *quality* of use. When the smoothing line crosses the midway point between the starting point and end point (approximately -1), that means speakers at that time produce tokens that are themselves phonetically intermediate between the end points. In Figure 2, however, when the smoothing line crosses 0.5, speakers don’t produce tokens where the verb is raised halfway to Tense.

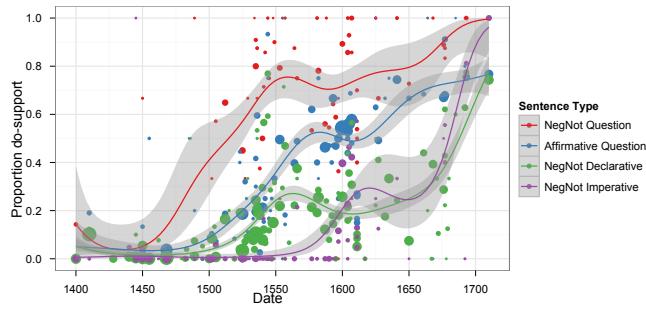


Figure 2: Rise of periphrastic *do* in Middle English.

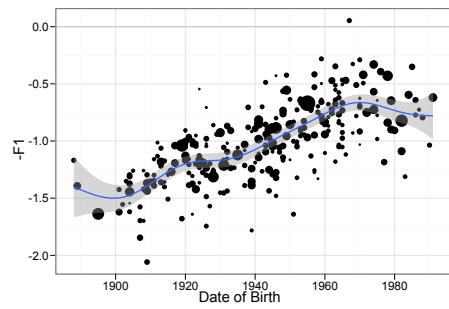


Figure 3: Raising of /ay/.

This has been, at least, the way which phonetic change has been understood within the field of sociolinguistics. Some may even consider the possibility that phonetic change is competitive change to be so implausible that an attempt to demonstrate it is not amounts to a straw man argument. However, the theoretical benefits of treating phonetic change as competitive are great enough that competitive phonetic change is worthy of falsification.

In fact, by the very fact that no one has tried to falsify competitive phonetic change, we *don’t* know that it isn’t true. In Figure 3, each point represents the mean F1 of a single speaker. The points appear to move continuously upwards across the time domain, but this is not a demonstration that phonetic change must be continuous. The same observation could be made of Figure 2. The points (representing texts) move continuously upwards over the time domain, but we *know* that the usage *within* authors was categorical, because a partially *do*-supported sentence is impossible. It is possible that the same is true in Figure 3. Each speaker may be producing tokens variably from two, categorical

distributions, and the frequency with which they draw from the innovative distribution increases over time. We then, as analysts, erroneously average over the productions of the two distributions, producing a view of apparently continuous change. Figure 4 illustrates this process. Each panel contains a different hypothetical speaker's data. In the farthest left panel, the speaker uses Category A 95% of the time, and Category B 5% of the time. Moving left to right, the hypothetical speakers use Category B to a greater degree, so that the speaker at the farthest right uses Category B 95% of the time. A speaker's categorical intentions to produce Categories A and B will, naturally, form two phonetic distributions, which are represented by the colored distributions in Figure 4. A researcher, presented with what is underlyingly two different distributions might mistakenly treat all the data as if it came from one, which is represented in Figure 4 by the grey distribution.

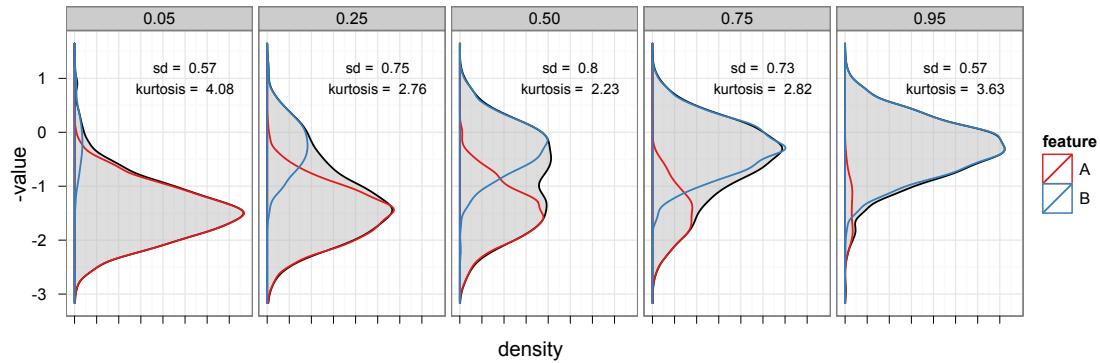


Figure 4: Phonetic change as changing frequency of two distributions.

This researcher could then calculate a few different descriptive statistics based on the undifferentiated distribution, specifically the mean, standard deviation, and kurtosis. Kurtosis is approximately a metric describing the relationship between the breadth of a distribution's peak to the length of its tails. If a distribution has relatively short tails relative to its peak, it has a low kurtosis value, and as the tails get longer relative to the peak, kurtosis goes up. A normal distribution, regardless of its mean and standard deviation, has a kurtosis of 3. In Figure 4, the calculated mean of the undifferentiated distribution increases from left to right. The standard deviation and kurtosis of the undifferentiated distributions also vary systematically as the mixture of Categories A and B change. Specifically, as the mixture of A and B approaches equality, the undifferentiated distribution should become wider, as indicated by the increasing standard deviation. Additionally, as the mixture approaches equality, the undifferentiated distribution should have shorter tails relative to the peak, as indicated by the decreasing kurtosis.

If we were to assume that the phonetic raising of pre-voiceless /ay/ in Figure 3 were due to changing mixtures of two, categorical distributions, then we should also expect to see a dip then rise in kurtosis, and a rise then fall in standard deviation. For the rest of this section, I will be comparing the empirical data from the PNC to these expectations by using simulation to see if this phonetic change is modelable as categorical competition.

First, in order to simulate the mixture of two distributions, I need to establish approximately what the mean and standard deviations of the two, differentiated distributions are. This is a non-trivial task because the whole point of this discussion is that researchers have *not* treated this phonetic change as progressing as categorical competition, and we have no principled basis upon which to state that the lower distribution ought to have one particular mean and standard deviation, while the higher distribution ought to have another. However, if phonetic change progressed as competition between two categories, we would expect speakers with the lowest over-all means to be using the new category the least, therefore the mean and standard deviation of their /ay0/ would be most likely to reflect the “true” mean and standard deviation of the low category. Additionally, we would expect speakers with the highest over-all means to use the higher category the most, therefore the mean and standard deviation of their /ay0/ would be most likely to reflect the “true” mean and standard deviation of the high category. I will be approximating the means and standard deviations of the two hypothetical categories, then, by estimating them based on the speakers with the highest and lowest over-all /ay0/ means. Figure 5 plots the relationship between the mean of F1 and the standard deviation of F1 for every speaker in the corpus. If these patterns were composed of two distributions, the lower distribution would have a mean of approximately 1.5, and standard deviation approximately 0.6. The higher distribution would have a mean of approximately 0.5, and a standard deviation of approximately 0.5.

For the simulation, I added to each of these 4 parameters ($\mu_a, \sigma_a^2, \mu_b, \sigma_b^2$), two more possible values, one slightly above, and one slightly below the values mentioned above, resulting in three possible values for each parameter. This creates the possibility of $3^4 = 81$ combinations of distributions to mix. Then, for 11 different mixture ratios (0%, 10%, 20%, … 100%) for these 81 possible distributions, I simulated 100 mixed distributions, and calculated their mean, standard deviation and kurtosis. Each of these 100 mixed distributions represent the values drawn from 1 speaker. So for each of the 81 mixtures, there are $11 \times 100 = 1100$ simulated speakers.

Figure 6 displays the data from one of the 81 mixtures ($\mu_a = 1.4, \sigma_a^2 = 0.6, \mu_b = 0.25, \sigma_b^2 = 0.4$). The standard deviation and the kurtosis are compared to the mean of each mixture. Standard deviations are bounded by 0 and ∞ , so I’ve plotted the log of the standard deviation. Kurtosis is bounded by 1 and ∞ , so I’ve plotted the log of the kurtosis minus 1. Normal distributions, no matter their variance, have the same kurtosis. I’ve indicated that value with a dashed line on the kurtosis panel.

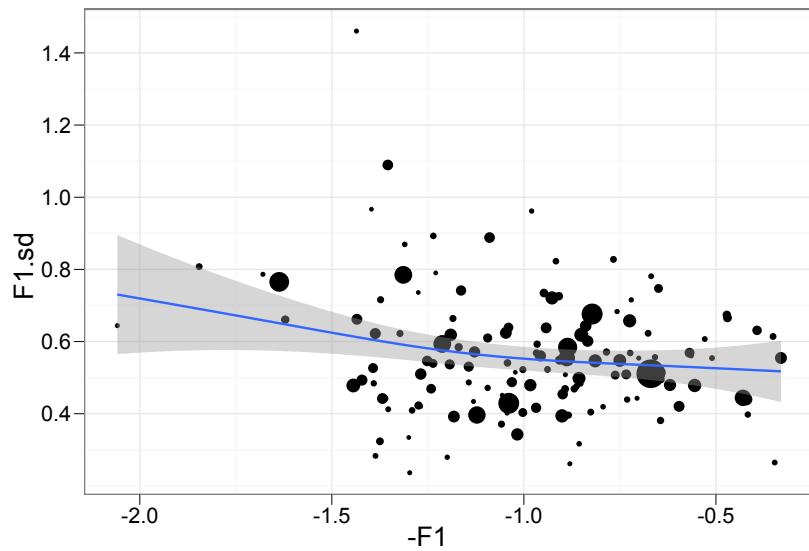


Figure 5: Relationship between mean $F1$ and standard deviation of $F1$ for /ay0/

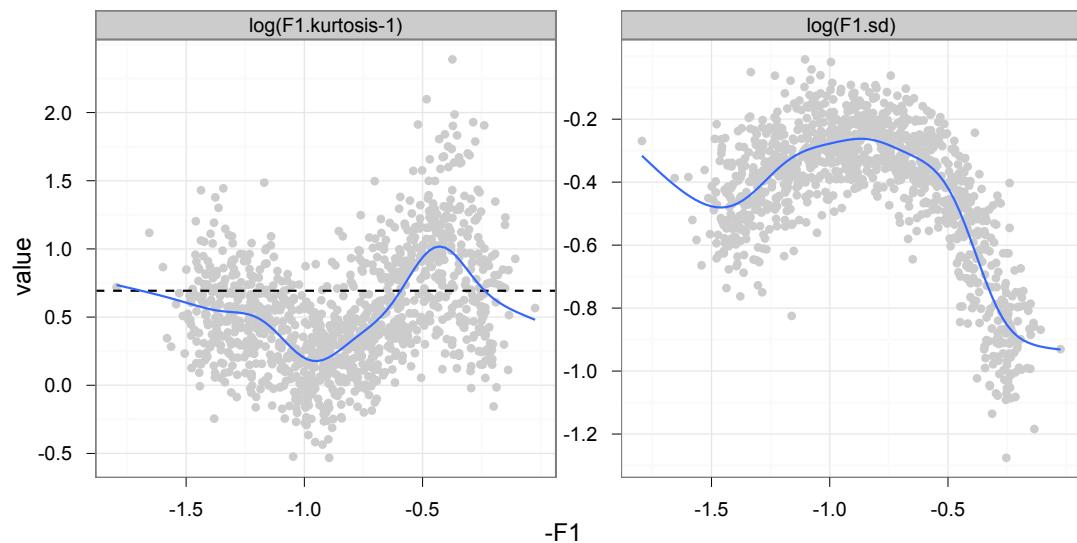


Figure 6: Results of one mixture.

As expected, kurtosis reaches a minimum near the midpoint of the change, and the standard deviation reaches a maximum. These change profiles can now be compared to the empirical change profile. This is done in Figures 7 and 8. Each facet is defined by the mean and standard deviation of the first distribution, meaning that there are 9 smoothing lines from simulations in each facet. The lines are colored according to the mean of the second distribution. The standard deviation of the second distribution isn't mapped to any graphical element. The solid black line represents the empirical relationship between the mean of F1 and the kurtosis F1 (in Figure 7) and the standard deviation of F1 (in Figure 8).

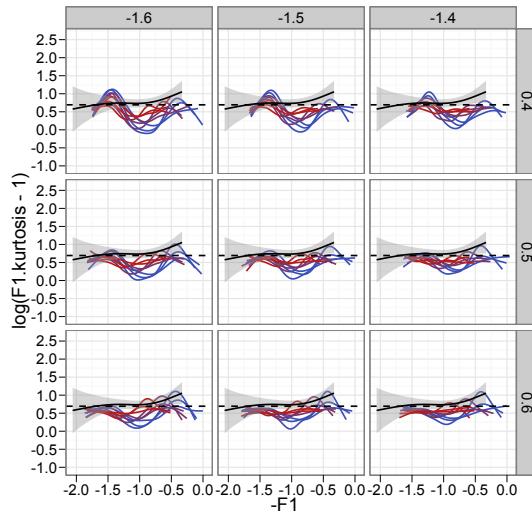


Figure 7: Comparison of F1 mean and F1 kurtosis

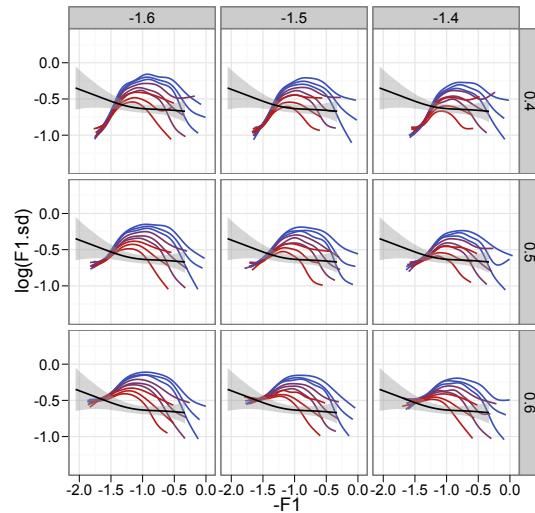


Figure 8: Comparison of F1 mean and F1 standard deviation

Looking at Figure 7, the only facet where the empirical line looks similar to the simulation lines is for $\mu_a = 1.4, \sigma_a^2 = 0.6$. This shouldn't be too surprising, since these parameter settings mean that the initial state began as closer to the final state, and was largely overlapping. That is, the less actual change was involved, the more the simulation resembled the empirical data. Another take away from Figure 7 is that while the very beginning and end of the change appear to move away from normality, the middle stage of the change actually exhibits the most normal-like kurtosis. This is, in fact, qualitatively opposite from what we would expect if the change progressed via different mixtures of two distributions.

There is no facet in Figure 8 that resembles the quantitative data. All of the simulations have maximum midway through their change, and the empirical data exhibits no

tendency to that effect. There does appear to be some kind of systematic relationship between the mean of F1 and the standard deviation, in that standard deviation appears to be decreasing as the mean increases. I don't have an explanation for this particular pattern, but we can be fairly certain that it is not due to the mixture of two distributions. In fact, when plotting these points across the time dimension, there does not appear to be any diachronic pattern of note *except* for the mean. Figure 9 illustrates.

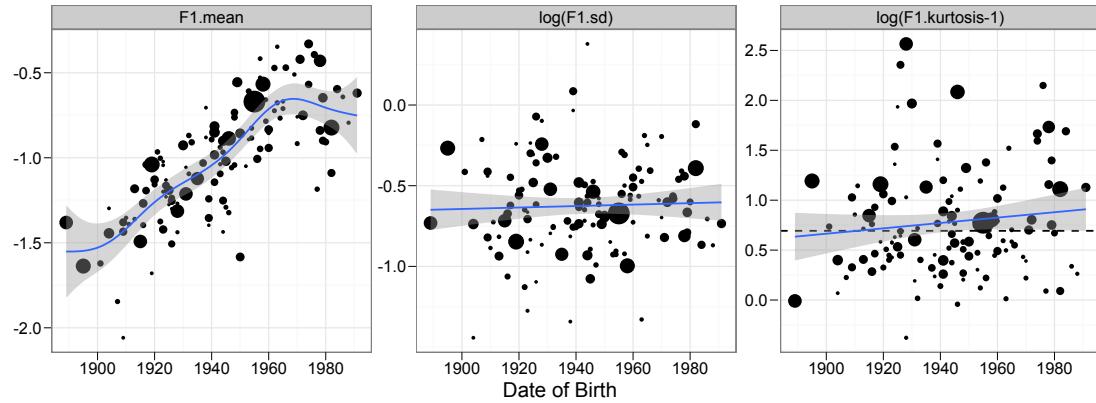


Figure 9: F1 mean, standard deviation and kurtosis across time.

While these simulations are based on quantitative analysis, the evaluation of their goodness of fit is entirely qualitative. I have already begun to model the competition model of phonetic change quantitatively, and will report on it in future work. The quantitative modeling involves a two layered regression. The underlying layer is a logistic regression which predicts the probability that a speaker would use distribution A or distribution B. The main predictor of this probability is date of birth, making the probability change over time. The second layer estimates the parameters (mean and standard deviation) of the distributions A and B. These parameters are time invariant. To fit this layered regression, I'm using Markov chain Monte Carlo estimation. So far, the results are inconclusive since the model does not appear to converge on stable estimates properly. This may be because the competition model is radically inappropriate for the data, or it may be due to my current unfamiliarity with MCMC estimation.

Once I have fine tuned the MCMC estimation, then it will be possible to quantify the model's goodness of fit in a number of ways. First, the model will produce estimates for the mean and standard deviations of the two competing distributions. The estimates that the current model produces make the means of the two distributions very similar, which does not fit the theoretical nor empirical data very well. So, the reasonableness of the over-all estimates can be evaluated. Next, the model will produce estimates for

every speaker. I can then, using the speaker level estimates, simulate new data from that particular speaker, and compare the speaker level simulation to the data that the speaker actually contributed. Finally, I'll be able to easily fit a model which only uses one distribution with a shifting mean, and compare its goodness of fit to the goodness of fit of the competition model.

There are some remaining adjustments I could make to both the simulation above and to the MCMC model. To begin with, the simulations above assumed that a given token only came from one distribution or the other. Perhaps instead, individual tokens are weighted averages of the two competing distributions. When going to speak, a speaker would then sample both distributions, and produce their weighted average. A process like this would be impossible within syntax, simply because of its modality. Within sound systems, however, halfway between two categorical distributions is well defined. This blending process would look like Figure 10. A weighted blend of two normal distributions is also a normal distribution, so the kurtosis should remain stable. However, the standard deviation actually decreases the more even the mix.

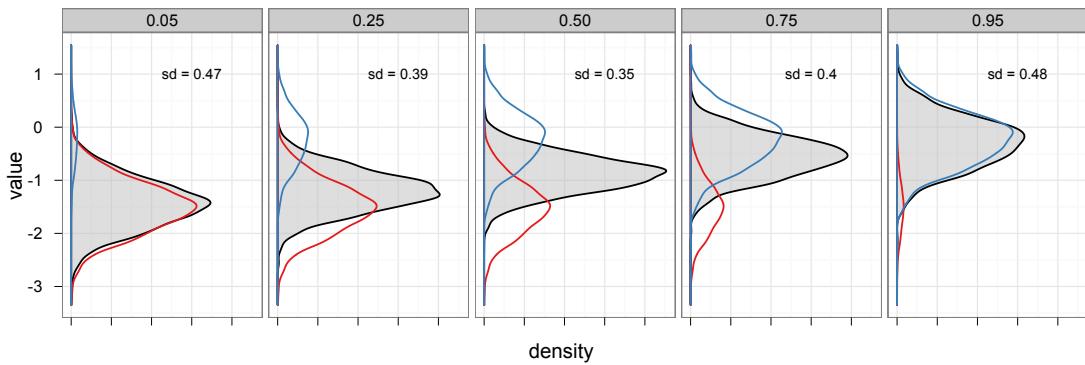


Figure 10: Phonetic change as weighted blend of two distributions.

I repeated the simulation from above, this time blending tokens rather than mixing. Again, the simulation lines don't match the empirical data well in Figure 11. They all produce a dip in the standard deviation around the midpoint in the change which is qualitatively lacking in empirical data, both when we compare it to the mean, and when we plot it across time.

Furthermore, competitive change plus blending actually creates the additional problem of how to deal with non-blending competition in phonology. Ideally, the blending of competitors would be made possible simply by the domain in which vowels exist. However, vowels can also vary without blending. A marginal example is presented in Figure 12, which plots every token from the PNC corpus of the first vowel in "either."

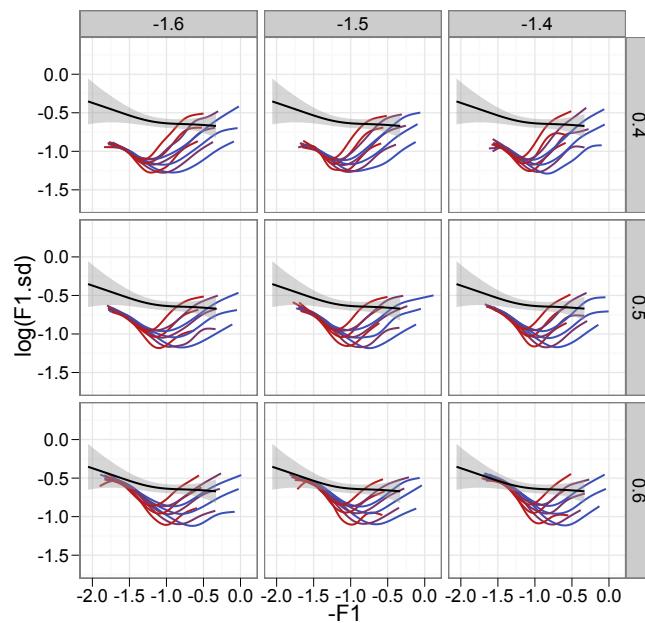


Figure 11: Comparison of F1 mean
and F1 standard deviation of blended distributions.

For Philadelphians, the first vowel is usually /iy/, but /ay/ is also occasionally used. Systematically absent from Figure 12 is a production like [eiðə̯]. Non-lexically based variation has also been found to occur without blending. For example, Brody (2009) found that while young children in Philadelphia varied in whether /æ/ was tense or lax in open syllables like “flannel,” the variation was not phonetically gradient. Rather, there was categorical variation between tense forms and lax forms.

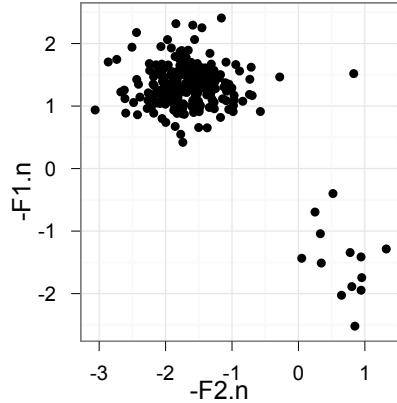


Figure 12: Variation between /i:ðə̯/ and /aɪðə̯/.

Another possible extension to the model I’ve presented above would be the addition of more distributions. Rather than the mixture or blending of two distributions, the change could be modeled as the mixture or blending of three or four distributions. However, I will not attempt to produce simulations of this sort. To begin with, addition of even a third distribution is theoretically unmotivated. The raising of /ay0/ moves from a low position to just about, but not quite, mid position. Most phonological feature theories only allow for one degree of height within this range, defined by [\pm low]. For every additional distribution we add to the simulation, we are also adding another possible degree of phonological height within this range, the only motivation for which is the need to fit phonetic change with a competition model.

Furthermore, for every additional distribution we add, we also add an additional, perhaps independent, phonetic change. By modeling phonetic change with two distributions, we are assigning each distribution to a meaningful location in terms of the change: the beginning and the end. There is no third meaningful location with regards to the change. We *could* place the third distribution midway between the beginning and the end. However, this drastically changes the nature of the phenomenon we are observing. Rather than a shift from A to B, we are observing two shifts: A to B (where B is the midpoint), then B to C (where C is the end). Determining the driving forces for a shift

from one beginning to one end is already complicated. The driving forces behind two consecutive, and uninterrupted shifts would be even more complicated.

It seems, then, that an attempt to model phonetic change as competition by adding additional distributions to mix is a case of the tail wagging the dog. Neither the data nor theory immediately lends themselves to such an analysis, and it would ultimately complexify the description of the phenomenon in an unjustified way.

Conclusions

The purpose of this section is to demonstrate how phonetic change is unique among language changes in that it progresses as a continuous movement through phonetic space, rather than as discrete variation between conservative and innovative forms. This fundamental fact about phonetic change is complicated by the fact that discrete, categorical phonology clearly plays an important mediating role, which will be discussed next.

3 Phonological Mediation

3.1 Early

The case study supporting the early influence of phonology on phonetic change which I will focus on most closely here is /ay/ raising.

The raising of /ay/ before voiceless segments was identified as a new and vigorous change in the 1970s (Labov, 2001). The next earliest report on the status of /ay/ in Philadelphia was Tucker (1944), who reported that Philadelphians made no difference in the quality of /ay/ between pre-voiced and pre-voiceless contexts. The age of the speakers who Tucker based this statement on is entirely up to speculation. However, loosely assuming the youngest possible age of his speakers would be in their late teens, and the oldest in their late 80s or 90s, the date of birth of his informants would have been somewhere between the 1860s and 1920s.

The raising of /ay/ is unique, because it has been argued to have phonetically natural origins (Joos, 1942; Chambers, 1973; Moreton and Thomas, 2004), but in almost all synchronic descriptions it exhibits opacity with regards to flapping, meaning that whatever natural phonetic properties of the context which favor raising there may be, they are eliminated by the flapping process. The key question which arises here is how /ay/-raising could apply in phonetically unnatural contexts if it began as a phonetically natural process.

Furthermore, at the current time in Philadelphia, the raised /ay/ allophone has undergone lexical diffusion (Fruehwald, 2007). The lexical items to most consistently exhibit exceptional raising are *spider* and *cider*. Most exceptionally raised words fit in a /aydV/

	Raised	Low
Surface True	rʌɪt (<i>write</i>)	raɪd (<i>ride</i>)
Opaque	rʌɪrə (writer)	raɪrə (rider)

Table 1: /ay/ raising in both surface true (i.e. phonetically natural) and opaque (i.e. phonetically unmotivated) environments.

frame, however, examples like raising in *tiny* and *cyber-* suggest that reanalysis of the surface flap as underlying /t/ is not tenable. All examples are monomorphemic, and fit in a /ayC \bar{V} / frame.

There exists, as I've called it, a consensus view as to how this movement from natural phonetic conditioning to abstract phonological conditioning takes place. In the terms of Ohala (1981), the raising of /ay/ before voiceless segments would be the result of "hypocorrection." At an early stage in the dialect, due to undershoot for some natural reason, a speaker might produce a slightly raised nucleus for /ay/ before a voiceless segment. A listener could then undo the coarticulatory effect, to recover the same underlying form. This correction for coarticulation is schematized in Figure 13. However, a listener might also fail to undo the coarticulatory effect, and recover a different underlying form. This hypocorrection is schematized in Figure 14.



Figure 13: Listener's correction of coarticulatory effects.



Figure 14: Listener's hypocorrection of coarticulatory effects.

It should be noted that this is only a plausible model if surface phonetic productions "sound like" phonological representations; that is, if the mapping between phonetic percepts and phonological representations is trivial and automatic. In theoretical frameworks where phonological representations are actually just stored phonetic mem-

ories, like Exemplar Theory, or where there is a strict and universal mapping between phonetic realizations and phonological representations (like Hale and Reiss (2008) prescribe), then the hypocorrection model works without further alterations. However, if phonetic realizations are language specific (e.g. Disner (1978) for vowels), and especially if phonological features purely abstract and contrastive representations (e.g. Mielke (2008)), then there is no sense in which a particular phonetic realization sounds like a particular phonological representation independently from the facts of the language at hand.

There are then two hypotheses for how this hypocorrection can play out when extended over time, depending on the theory of grammar one chooses. Bermúdez-Otero (2007) outlines the view from a modular-feedforward grammatical model. First, a language specific implementation rule gets added to the grammar. Then, as the magnitude of the difference in phonetic realizations between voiced and voiceless contexts grows, it may be reanalyzed as a phonological process.

An alternative hypothesis comes from Exemplar Theory, where phonetic experiences are stored in detail. In this model, productions of /ayt/ which were heavily affected by coarticulatory pressures are re-stored in detail. Through multiple cycles of the production-perception loop, the /ay/ category undergoes a slow creep towards higher and higher productions before voiceless segments. Pierrehumbert (2001) has shown this to be a conceptual possibility through simulation. One apparent drawback of the model is that in its simplest form, it predicts that the category should slowly increase in its variance over time, since both low and high exemplars are now stored in the exemplar cloud, where originally there were only low exemplars. As I illustrated above, this does not happen (see Figure 5, repeated here as Figure 15). Wedel (2006) demonstrated, also through simulation, that if the method of production in exemplar models sampled multiple exemplars and averaged them, that the variance of the category would *not* increase over time. However, it is not immediately clear how this averaging procedure would avoid eliminating phonological variation within a word, like that in *either*, also mentioned above. Presumably some additional principle would have to prevent both [i:ðə] and [aɪðə] tokens from being sampled, because their average ([eɪðə]) would be produced, which would eventually eliminate either the [i:ðə] or [aɪðə] pronunciation.

Recently, Baker et al. (2011) raised some objections to the consensus view laid out above. They pointed out that the actuation problem, as formulated by Weinreich et al. (1968), not only requires an explanation for why the language changes which happened did happen, but also for why all possible language changes which could happen (including highly probable ones) *didn't* happen at any given point in time. The *accumulation of error* model, as they call it, makes the raising of /ay/ an inevitability, rather than a possibility. They propose, instead, that before a change occurs, there is a broad range of variation across individuals in a speech community in the phonetic production of

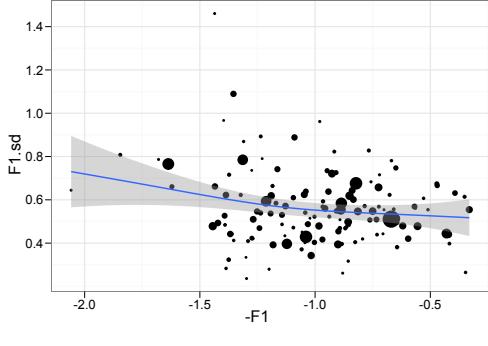


Figure 15: Relationship between mean F1 and standard deviation of F1 for /ay0/

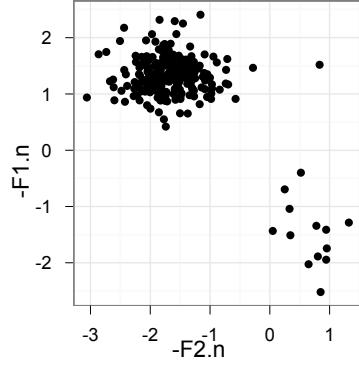


Figure 16: Variation between /i:ðə/ and /aɪðə/.

some segment under coarticulatory pressure. They support their model with empirical, synchronic data on /s/ retraction in initial /str/ clusters. The phonetically natural inter-speaker variation provides the seeds for phonetic change. Then, a rare alignment of speakers who have a large degree of /s/ retraction *and* social influence occurs, setting off the language change.

While the data on /s/ retraction appears to support their model, the data on /ay/ raising does not. The initial state of the speech community before the change, according to Baker et al. (2011), should have a high degree of variance, with speakers covering the whole range from entirely low /ay/ to entirely raised /ay/. The mechanism of change would then be the gradual settling of the speech community on one target or the other. This mechanism should have many similarities to the discrete variation model which I demonstrated earlier is not a good fit for phonetic change, especially since they say that the initiation of the change should be phonetically abrupt. Figure 17 (which is Figure 3 reproduced) displays the raising of /ay/ over date of birth. The remarkable aspect of the data is how little variance there is at any given point in time, and hardly the kind of inter-speaker variation needed to seed the change in the manner of Baker et al. (2011). No speakers have a $-F1$ greater than -1 until about 1920, and very few speakers have a $-F1$ less than -1 after 1960. Furthermore, there doesn't appear to be anything particularly phonetically abrupt about the change, not along the time domain, between speakers, nor within speakers (as illustrated in §2).

Coming back to the view from a modular-feedforward model of grammar, it is an open question as to how long the phonetically natural conditioning due to hypocor-

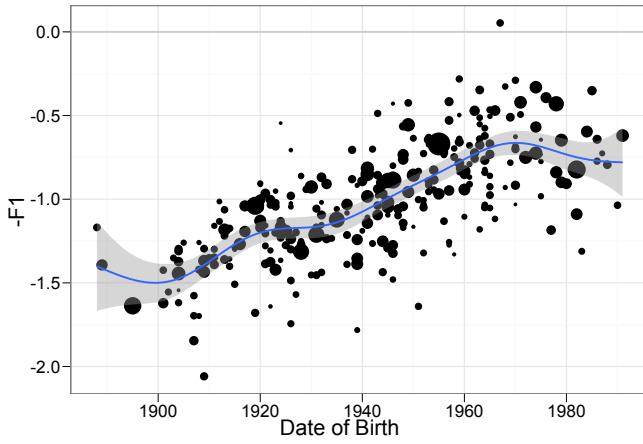


Figure 17: Raising of /ay/.

rection should persist before phonological reanalysis takes place. However, Janda and Joseph (2003) propose that it should be a relatively short period of time between a stage of pure phonetic conditioning and phonological generalization. However, it's not clear that the examples which they discuss are especially relevant to phonetic change. Despite appealing to pure phonetic conditioning, they frame all three of their examples in terms of discrete variation between two variants (Romance vowel prothesis: $\emptyset \sim e$; Swiss German vowel lowering: $o \sim \circ$; contemporary English s-retraction: $s \sim \emptyset$). Furthermore, they don't define "brief", nor how extreme the pure phonetic conditioning must be before it will be reanalyzed as phonological.

A synthesis of Janda and Joseph (2003) and Bermúdez-Otero (2007) would produce the following sequence of events in the shift of pre-voiceless /ay/:

1. Due to natural physiological and/or perceptual pressures, the phonetic realization of /ay/ before voiceless segments would be produced with slightly higher nuclei.
2. After a brief period of pure phonetic conditioning, the raising of /ay/ would be reanalyzed as a phrase level phonological rule.
3. As the /ay/-raising process gradually moves up into more abstract phonological domains, aspects of its conditioning may be altered within those domains.

The questions which we can answer empirically, using the Philadelphia Neighborhood Corpus, are the following:

- (a) Was there ever a period where /ay/-raising was purely phonetically conditioned?
How long did it last?
- (b) When did /ay/-raising begin occur opaquely, i.e. before flaps?
- (c) When did raised /ay/ begin to occur in lexically specified conditions?

Pure Phonetic Conditioning

One important diagnostic for pure phonetic conditioning of /ay/ raising is whether it occurs at the phrasal level. Following Janda and Joseph (2003); Bermúdez-Otero (2007), we might expect to see word final /ay/ undergoing raising when followed by a word with a voiceless segment. In fact, Idsardi (2007) reports intuitions for his Canadian English dialect that raising can apply *opaquely* in phrases like “Don’t lie to me” [lʌɪ.ɹə.mi]. According to Janda and Joseph (2003), however, this phrase level raising may disappear as the change is reanalyzed and reorganized phonologically.

To establish whether phrase level raising ever occurred in Philadelphia, I took a subset of the /ay/ data from the PNC with /ay/ in word-final position, followed by either a voiced or voiceless obstruent (following nasals, glides, /r/, /l/, pause, and other non-linguistic noises are excluded). For every speaker, I calculated their mean F1 in voiced and voiceless contexts, keeping following stressed and unstressed syllables separate. Following stress has been implicated in blocking raising in other dialects (Chambers, 1973), so it seems important to account for that here. Additionally, the phonetic pressures of a following stressed syllable versus an unstressed syllable are not likely to be identical.

Figure 18 displays the results. Each point represents a mean value in a voicing context for a speaker, scaled according to how many tokens the speaker had. The smoothing lines are cubic regression splines. It does not appear that there was ever any effect of the following voicing context on word-final /ay/. This impression from the graphical display is confirmed by careful statistical analysis.

For each of the stress contexts, I fit three generalized additive regression models, where the date of birth of the speaker was fit by a cubic regression spline. It was possible to include a term which weighted the spline according to voicing context, i.e. allowing the different voicing contexts to have different curves over time. I compared this to a model where both voicing contexts had the *same* curves over time, and then this to a model which included no information about the voicing context at all. In addition, random intercepts for speaker and for the /ay/ lexical item were included.

1. $F1 \sim \text{Voicing} + s(\text{DOB}, \text{by} = \text{Voicing})$
i.e. Different curves for different voicing contexts.

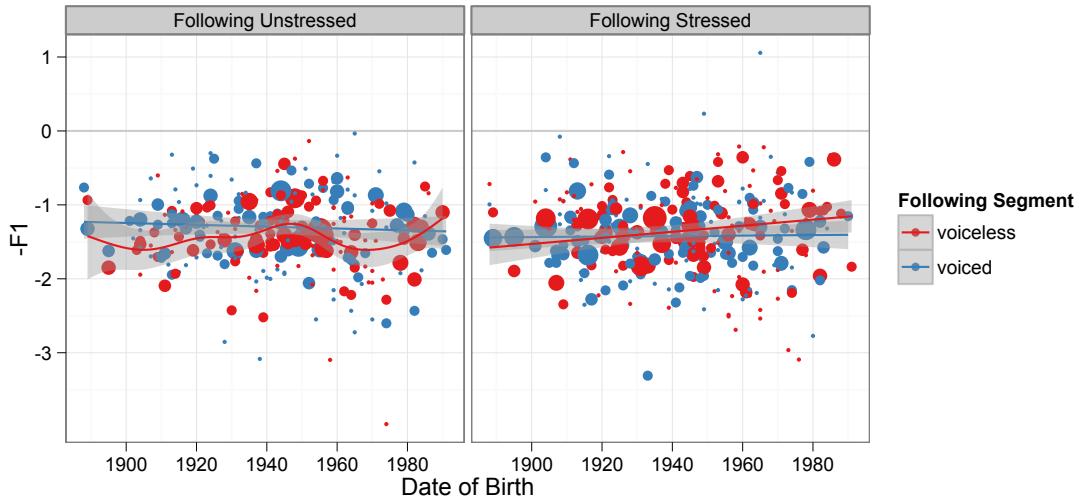


Figure 18: Word final /ay/, followed by following voiced and voiceless obstruents, faceted by the stress of the following syllable.

2. $F1 \sim \text{Voicing} + s(\text{DOB})$
i.e. Same curves, but transposed up or down by voicing context.
3. $F1 \sim s(\text{DOB})$
i.e. Same exact curves.

Table 2 displays model comparison metrics (AIC and BIC) for the goodness of fit of these three models for each stress context. The AIC and BIC are both information criteria that penalize the goodness of fit of a model by the number of parameters it uses. There is no absolute criterion for determining a significant difference in AIC or BIC, but over-all, smaller values mean more probable models. In both stress contexts, for both the AIC and BIC, the models where there is no voicing effect across word boundaries whatsoever are preferred.

Now, it is probably the case that not all word-final contexts are alike phonetically. The initial segment of the following word may exert variable influences on the preceding /ay/ depending on a number of prosodic factors. As a quick check, I extracted just those word final /ay/ tokens which were followed by *to* and *the*. These two words are, fortunately, the most frequent words to follow a word-final /ay/, and they are both also very weak elements, therefore most likely to lean on the preceding /ay/. They also,

Model	Following Stressed		Following Untressed	
	AIC	BIC	AIC	BIC
M1: Different Curves	8272.9	8304.6	6918.0	6948.5
M2: Same curves	8270.6	8297.8	6912.9	6939.1
M3: No voicing effect	8269.3	8291.9	6912.2	6933.9

Table 2: AIC and BIC for the three models of phrase level /ay/ raising.

very conveniently, contrast in the voicing of their initial segments. Figure 19 displays the mean values of /ay/ followed by *to* and *the* by speaker, over date of birth. As with the larger data set of word final /ay/, there does not appear to be any effect of the voicing of the following segment.

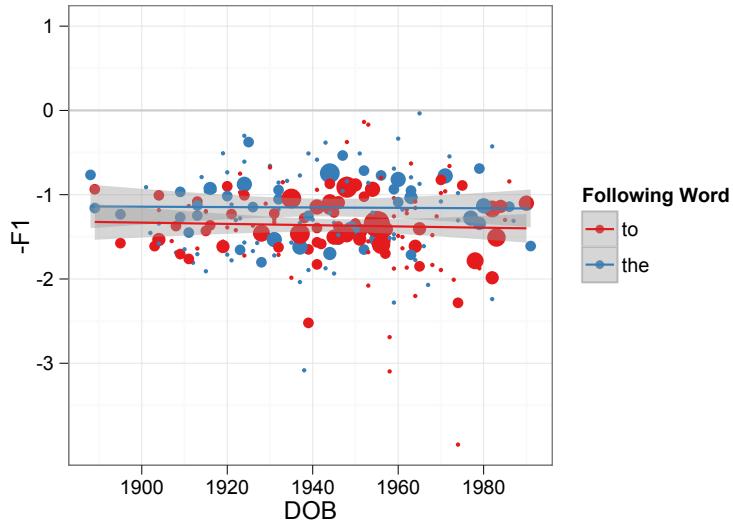


Figure 19: Word final /ay/, followed by following *to* and *the*.

In conclusion, there is no evidence that /ay/-raising applied across word boundaries at any time. It might be worth asking whether the relevant time period where /ay/-raising would have been strictly phonetically conditioned is contained within our data. It could be the case that a period of pure phonetic conditioning (which would allow for /ay/-raising to occur across word boundaries) preceded the time period in our data. However, this appears to be unlikely. In the next section, I'll be specifically analyzing

the size of the difference between /ay/ in different voicing contexts, and at the earliest time periods, there is only a marginal difference in height. Additionally, the earliest dates of birth in our data corresponds with the dates of birth of speakers who Tucker (1944) was most likely to be listening to when he said that Philadelphians made no distinction in vowel quality in /ay/ according to voicing context.

Opacity

It might be argued that the fact that /ay/ never underwent raising across word boundaries is not in itself evidence against a stage of pure phonetic conditioning. There are, perhaps, robust phonetic differences between /ayT/ and /ay#T/ (as was suggested in Janda and Joseph (2003)). Examining how /ay/-raising was conditioned in pre-flap contexts diachronically allows for much firmer conclusions. The pre-flap context is, in fact, an almost perfect object for analysis, because it can provide for us evidence of when phonological generalization occurred relative to the onset of the phonetic change.

To begin this analysis, I first had to decide upon the right subsets of the data to compare. First, I looked only /ay/ followed by a /t/ or /d/, in order to keep the flapping vs non-flapping contexts as similar as possible in terms of following place and manner. Flapping is not yet annotated in the PNC, and determining whether it has applied in any given instance (either by impressionistic coding, or by some automated classifier) will be left to future work. Instead, the flapping context was defined as follows.

(1) Flapping Context

- (a) /ay/ followed by /t/ or /d/
- (b) The /t/ or /d/ is then followed by an unstressed vowel.
- (c) The entire /aytV/ or /aydV/ sequence is contained within the same word.

This is an easily definable subset of the PNC. One feature which it does not take into account is the location of morpheme boundaries relative to any of the segments involved, which may be another property which will be added into the analysis in the future. Some specific lexical items were excluded from this subset, specifically those which were reported to have exceptional raising in Fruehwald (2007) (e.g. *spider*), and any where the /t/ might have been glottalized before a syllabic /n/ (e.g. *frighten*).

For a comparison subset, I chose /ay/ followed by a word-final /t/ or /d/. This is perhaps, the most archetypal, and least complicated context for raising to apply. The subset was defined this way.

(2) Non-Flapping Context

- (a) /ay/ followed by /t/ or /d/

(b) The /t/ or /d/ is then followed by a pause, annotated as “{sp}” in the PNC.

I specifically chose /t/ and /d/ followed by a pause in order to avoid any possible complications associated with word final /t/ or /d/ in connected speech. For example, I did not want to potentially include flaps which were formed across a word boundary, nor tokens where the /t/ and /d/ could have been strongly lenited. It is worth noting that, occasionally, the forced aligner mistakes a long closure and release as a pause. However, this is an acceptable form of error for my purposes, since regardless of whether the pause was “real,” the /t/ or /d/ was certainly not lenited nor flapped.

Table 3 displays the number of tokens for each segment by context.

	C#	internal flap
T	1134	305
D	598	349

Table 3: Number of Tokens by Segment and Context

For visual display, I calculated the mean F1 for each segment for each context for each speaker. Figure 20 displays the cubic regression splines based on these means.³ As can be seen, /ay/ appears to pattern diachronically along with the underlying voicing of the following segment, rather than with the surface voicing. That is, visually, it appears that as soon as there was any phonetic difference between /ayt/ and /ayd/, there was already a difference in the flapping contexts, meaning that /ay/-raising applied opaquely from its outset as a phonetic change.

In order to establish this quantitatively, for each segment (/t/ and /d/) I fit three GAMs using cubic regression splines using the raw data, including random intercepts for speaker and for word. The first model allows the diachronic smooths to vary by context (flapping vs. final-C), and the second did not. The third model didn’t include a term for the context at all. The AIC and BIC criteria for model comparison are given in Table 4.

Again, there is no absolute standard by which to determine a significant difference in AIC or BIC, but the smaller, the better. With the exception of the AIC for the /d/ models, the models where there is no effect of flapping context at all are preferred. Figure 21 displays the model fits for /t/ and /d/. For /t/, I’ve plotted the model with no flapping effect at all, since that was preferred. For /d/, I’ve plotted the model

³The GAMs in these figures were weighted by $\log(n + 1)$ where n is the number of tokens from a speaker. I used $\log(n + 1)$ rather than $\log(n)$ so that speakers who had one token would not be given a $\log(1) = 0$ weight.

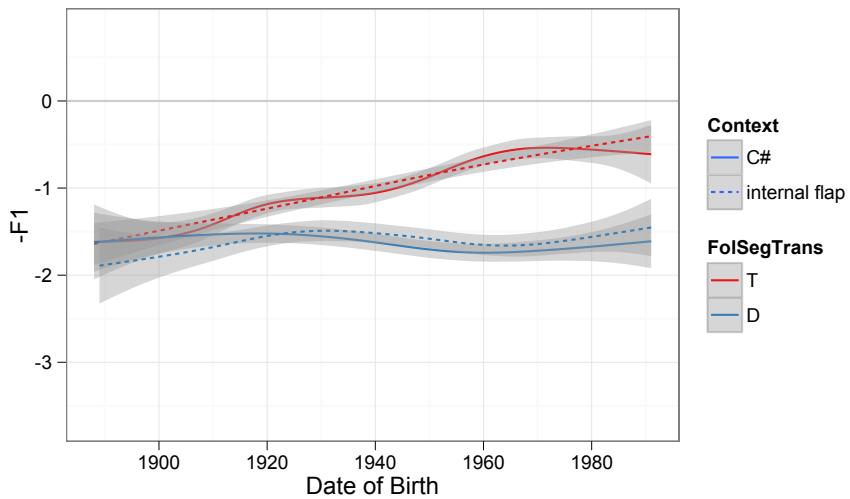


Figure 20: /ay/-raising in flapping and non-flapping context

Model	/t/		/d/	
	AIC	BIC	AIC	BIC
M1: Different curves	2475.5	2512.4	1484.3	1518.3
M2: Same curves	2467.7	2499.3	1485.6	1514.7
M3: No context effect	2465.7	2492.0	1485.1	1509.4

Table 4: AIC and BIC for three models /ay/-raising, comparing flapping contexts to non-flapping contexts.

allowing different curves. Even though the BIC preferred the model with no context effect by a larger margin than the AIC preferred the different curves model, I've plotted the different curves model just to see how different the curves are actually predicted to be.

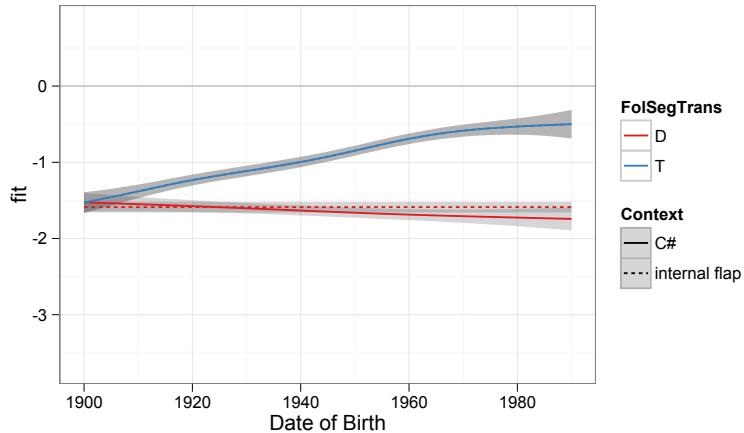


Figure 21: /ay/ raising and flaps GAMM model fits

Even the model which allowed for maximal difference between the flapping and non-flapping context only predicts a very marginal diachronic pattern for /d/.

These results make it appear that /ay/-raising was conditioned by an abstract phonological context from the very outset. Regardless of the surface context, /ay/ patterns together diachronically based on the underlying phonological voicing of the following segment. A possible objection to this conclusion is that the purely phonetic effect of a /t/ might persist when it is flapped, so-called incomplete neutralization. It is worth, at this point, to describe the two primary hypotheses about the phonetic naturalness of /ay/ raising.

Perhaps the most familiar hypothesis of the phonetic origins of /ay/-raising is the effect of pre-voiceless shortening, a well known cross-linguistic phenomenon (House and Fairbanks, 1953; Chen, 1970; Tauberer, 2010). Joos (1942) and Chambers (1973) both appeal to pre-voiceless shortening as the source of phonetic naturalness for Canadian Raising. The argument goes that since /ay/ has a shorter duration within which to make a gesture from low-back to high-front, speakers undershoot the nucleus, leading to assimilation of the nucleus to the glide.

If pre-voiceless shortening provides the phonetic grounding for /ay/-raising, then for /ay/-raising before flaps to be phonetically natural the duration difference between underlyingly voiced and voiceless flaps would have to persist in the surface phonetics.

Braver (2011) argues that vowels before flapped /t/ and /d/ *do* differ in duration, but the size of the difference is very small (8.76 ms). I would argue that in order for the pre-voiceless shortening effect to be the pure phonetic conditioning for /ay/-raising, the duration difference between /ay/ before flapped /t/ vs flapped /d/ must be commensurate with the degree of height difference.

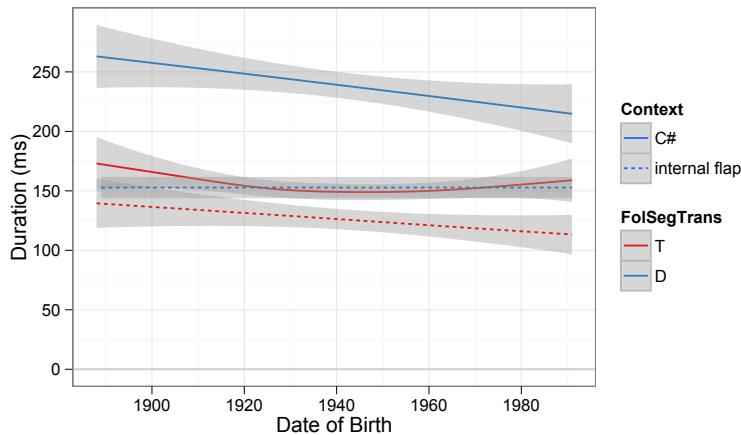


Figure 22: /ay/ Duration

Figure 22 displays diachronic patterns of /ay/ duration in the four contexts of interest. The cubic regression splines are based on the median duration of /ay/ in each context by speaker. There does appear to be a sizable difference in duration between /ay/ before flapped /t/ and flapped /d/. A non-parametric estimate of the differences⁴ (the pseudo-median) is approximately 32 ms, which is a significant difference according to a Wilcoxon Signed Rank test ($p = 4 \times 10^{-7}$). However, the size of this difference before flaps is less than half that of the difference between surface /t/ and /d/ (pseudo-median = 80 ms), and /ay/ before flapped /t/ is not half as raised as /ay/ before surface /t/, it is raised to exactly the same amount.⁵

Moreover, it is not the relative difference in duration between underlying voiced and voiceless segments which is relevant to the phonetic conditioning. Rather, it is the absolute duration. That is, if surface /t/ triggers raising because of the shortening effect it has on /ay/, then other contexts which have similar shortening effects on /ay/ should also trigger raising, which according to Figure 22 would also include flapped /d/. However, /ay/ before flapped /d/ has never been subject to raising, according to the results above. In conclusion, the contexts under which /ay/ undergoes raising does

⁴This was calculated over (d-flap – t-flap) within speakers.

⁵Wilcoxon Signed Rank Test on pairwise differences between surface /t/ and flapped /t/: $p = 0.79$

not appear to be explained by pure phonetic duration.

I should, however, mention a shortcoming of the available duration data. The FAVE suite was specifically designed to not attempt to measure vowels shorter than 50 ms, meaning that the current available duration data is censored. This could have an effect on some of the duration results, so I will have to collect more complete duration data from the PNC to confirm these results.

An alternative hypothesis of the phonetic origins of /ay/-raising off-glide peripheralization, as given by Moreton and Thomas (2004). Moreton and Thomas (2004) point out that in dialects which have /ay/-monophthongization, the pre-voiceless context resists monophthongization the most, which would be surprising if pre-voiceless contexts had the most phonetic pressure to reduce the size of the diphthong gesture. They suggest that instead, the pre-voiceless context peripheralizes the glide in /ay/, creating pressure for the nucleus to assimilate. Moreton (2004) argues that this is a natural, cross-linguistic phenomenon.

Unfortunately, we don't yet have glide measurements for the PNC, but they are something that I propose to add. However, Rosenfelder (2005) collected glide measurements for her study of Canadian raising in British Columbia. Figure 23 plots the glide trajectories for /ay/ in pre-voiced, pre-voiceless, and pre-/t/-flap contexts.⁶ Again, in terms of offglide peripherality, /ay/ before /t/-flaps doesn't appear to pattern with surface /t/. Rather, its offglide peripherality looks more similar to surface voiced segments.

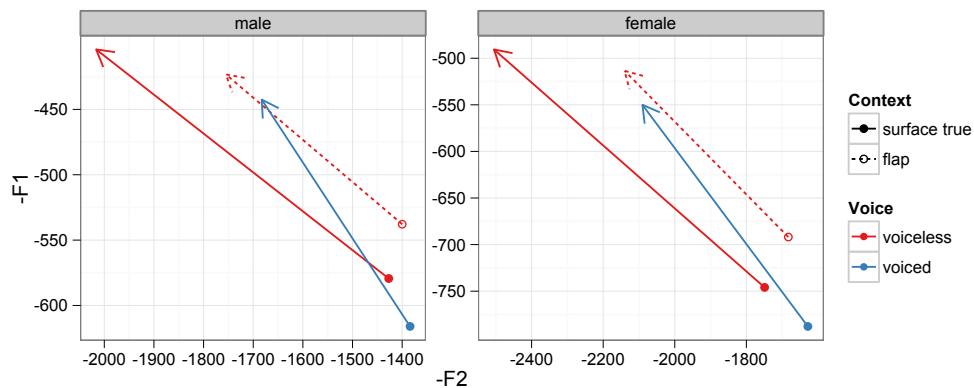


Figure 23: /ay/ trajectories in British Columbia (Rosenfelder, 2005)

Figure 24 summarizes how the /ay/ contexts under discussion pattern together into phonetically natural classes, defined by the effect they have on /ay/. When it comes

⁶Data for Figure 23 was adapted from Figures 4.1 and 4.2 in Rosenfelder (2005, p 53–54) using Plot Digitizer (Huwaldt, 2011).

to duration, surface flaps, regardless of underlying voicing, and surface /t/ pattern together. For offglide peripherality, surface flaps appear to pattern with surface /d/. Only when it comes to /ay/ height do flaps pattern according to their underlying voicing.

Duration	Glide Peripherality	Height
$\left\{ \begin{array}{l} t \\ r^- \\ r^+ \\ \{ d \} \end{array} \right\}$	$\left\{ \begin{array}{l} t \\ r^- \\ (r^+) \\ d \end{array} \right\}$	$\left\{ \begin{array}{l} t \\ r^- \\ r^+ \\ \{ d \} \end{array} \right\}$

Figure 24: Phonetic Natural Classes

It truly appears as if /ay/ raising was triggered by the underlying phonological voicing of the following segment, rather than the surface phonetic properties of its context. The consequences of these results are wide-reaching. First, it suggests that the consensus view of conditioned sound change (error-accumulation) cannot be accurate. There is no identifiable period where the conditioning on /ay/-raising was purely phonetic. The conditioning on /ay/-raising can be compared to conditioning that *is* purely phonetic, the effect of a following nasal on /aw/, given in Figure 25. The raising of /aw/ along the front diagonal of the vowel space is another one of the new and vigorous changes identified in the 1970's in Philadelphia (Labov, 2001). In Figure 25, this diagonal movement is represented on the y-axis. The effect of a following nasal in advancing this raising process has been noted. Every speaker in the PNC is represented in Figure 25 by 2 points. The green point represents their mean of /aw/ in pre-nasal contexts, and the purple point represents their mean of /aw/ in all other contexts.

We can tell that the effect of a following nasal on /aw/ is purely phonetic because the diachronic patterns for pre-nasal /aw/ is strictly parallel to the diachronic pattern of /aw/ in other contexts. This parallelism is in sharp contrast to the pattern of pre-voiceless /ay/, which as I have shown above, changes rapidly and separately from /ay/ in all other contexts. Under the error accumulation model, the conditioning of /ay/-raising should have originally been similar to the conditioning of /aw/ raising, but based on the above results it never was. Instead, phonetic conditioning like that of nasals on /aw/ is qualitatively different from the phonological conditioning of voiceless segments on /ay/.

However, even the models proposed by critics of the error accumulation model don't appear to be supported by the data. There does not appear to be even a brief period of pure phonetic conditioning, per the Big Bang model (Janda and Joseph, 2003). Nor is the onset of the change phonetically abrupt, in the sense that it starts out as a large phonetic

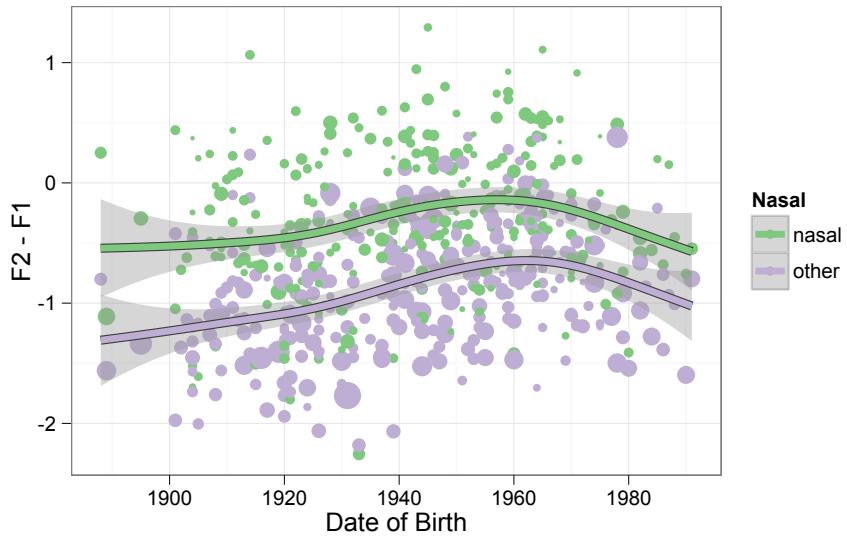


Figure 25: Phonetic conditioning of /aw/-raising

difference, per Baker et al. (2011).

From these results, I have to conclude that the phonological conditioning of /ay/-raising entered the dialect either simultaneously with the phonetic change, or preceded the phonetic change. This constitutes my answer to the question “When does phonological reanalysis take place?” Rather than suggesting it takes place towards the end of the change, midway, or even very early on, these results suggest that it takes place at the very onset. In early stages of the change, there must have been an abstract phonological difference between /ay[–voice]/ and /ay[+voice]/ that corresponded to only a minimal phonetic difference. At first glance, this may appear to be a strongly counterintuitive result, but the possibility of categorical phonological distinctions corresponding to small phonetic differences is well known to occur in the case of near-mergers (Labov et al., 1972; Labov, 1994), including the *ferry-furry* near-merger in Philadelphia.

These results also contradict the notion that new phonological rules are necessarily added to the “end” of the grammar (King, 1973; Bermúdez-Otero, 2007). In a simple rule-ordering approach, it appears that /ay/-raising always applied opaquely, before flapping. In a stratal approach, it appears that /ay/-raising enters the grammar at the word level, both because it applies opaquely relative to phrase level flapping, and it never applied across word boundaries.

Analogy?

An alternative hypothesis to phonological conditioning might be analogy. Most of the /ay/'s in /t/-flap contexts are morphologically complex words (e.g. *fight-ing*). It could be argued from an Exemplar Theoretic standpoint that *fighting* has a raised /ay/ by analogy to *fight*. One prediction this argument makes that is different from phonological conditioning is that roots which appear in flapping contexts more often than they do in non-flapping contexts should be less likely to undergo raising, or at least they should do so at a slower rate.

The PNC has not yet been morphologically parsed, but as an expedient, I instead examined different /CayC/ frames. First, I looked at /Cayt/ sequences, and calculated the ratio of how often they occurred in non-flapping contexts to flapping contexts. For example, the sequence /fayt/ appeared in 177 non-flapping contexts (all *fight*), and in 70 flapping contexts (all inflections or derivations of *fight*, like *fighters*), for a ratio of 2.5. I then examined how these ratios affected the rate of raising in flapping contexts. The hypothesis would be the higher the ratio, the more often /ay/ occurs in a transparent context in the frame, the more likely /ay/-raising is to have analogized to the opaque context for this frame. Figure 26 displays the means per-speaker, binning the ratio of transparent-to-opaque, for /t/-flaps. The “3+ x” bin are /Cayt/ frames which occur most often in transparent environments, and the “0 to 1x” bin contains frames which occur most often in flapping environments. There is essentially no effect of the rate at which a frame occurs in a transparent context and the degree to which it raises, at least not one which is consistent with the Exemplar Theory hypothesis.

I carried out another analysis, this time looking strictly at flapping contexts. For each [Cayr] frame, I calculated the frequency with which the flap was underlyingly /t/ or /d/, and calculated the /t/-to-/d/ ratio. For example, for the sequence [sayr], the flap corresponded to a /d/ 111 times (mostly *decided*), and to a /t/ 34 times (mostly *excited* and *exciting*) for a /t/-to-/d/ ratio of 0.30. I then examined the effect of the /t/-do-/d/ ratio on the rate of raising before underlyingly /t/ flaps. The hypothesis is that the higher the ratio, the more often the flap corresponds to an underlying /t/ for the frame, the more likely /ay/-raising is to occur in the frame. Figure 27 displays the means per-speaker, binning the /t/-to-/d/ ratio. There is still essentially no effect, although the “3+ x” frames appear to be somewhat lower than the rest, exactly the opposite of the hypothesis.

The results of this preliminary examination seems to indicate that analogy does not appear to play a major role in the opacity of /ay/-raising, and does not provide a better explanatory account than the phonological conditioning account provided here.

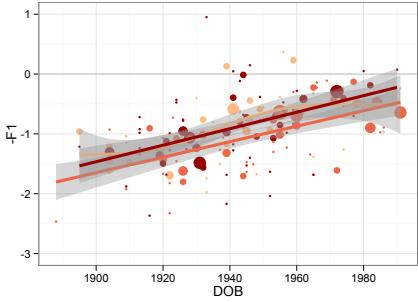


Figure 26: /ay/ raising before /t/-flaps, by the ratio of transparent-to-opaque frequency for the given /Cayt/ frame.

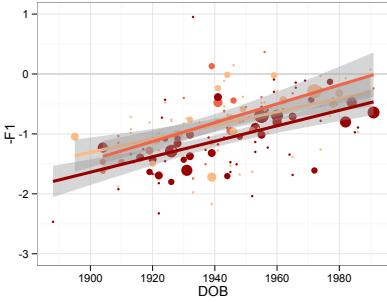


Figure 27: /ay/ raising before /t/-flaps, by the ratio of /t/-to-/d/ frequency for the given [Cayr] frame.

Lexical Diffusion

If /ay/-raising applied opaquely from its very onset, the next question to answer is when it began to be lexically specified in items like *spider*, *cider* and *Snyder*. According to most hypotheses, the lexical diffusion of raised /ay/ should occur some time after the onset of opacity, which we now know is synchronous with the onset of the phonetic change. Unfortunately, words exhibiting exceptional raising weren't added to the standard word lists until 2006 (following preliminary results eventually discussed in Fruehwald (2007)), and they are not exceptionally frequent in free conversation. In the entire PNC, there are a total of 62 tokens of words which have been documented to definitively have exceptional raising (52 *Snyder*, 10 *spider*) from 31 speakers. Since the numbers are so low, and the speakers so few, I will rely primarily on visual inspection of graphical displays of the information in this section, rather than heavy duty statistics.

My first step in the analysis was to calculate the mean height of the exceptionally raised words by speaker, which will call the SNYDER class, because tokens of *Snyder* outnumber *spider* 5-to-1. I then also calculated the mean height of /ay/ before /t/ and other /d/ words for those same speakers. Figure 28 displays the smoothing lines for /ayt/, /ayd/, and the SNYDER class for these speakers. It appears that on average, the *Snyder* class was low until approximately 1940. Then it began to pattern with the raised allophone, a change which was completed by 1960.

The change which took place in the *Snyder* class can be compared to a lexical item which did not undergo the change *Friday*. Why *Friday* has not undergone exceptional raising is unimportant at the moment, but it may be due either to its status as a com-

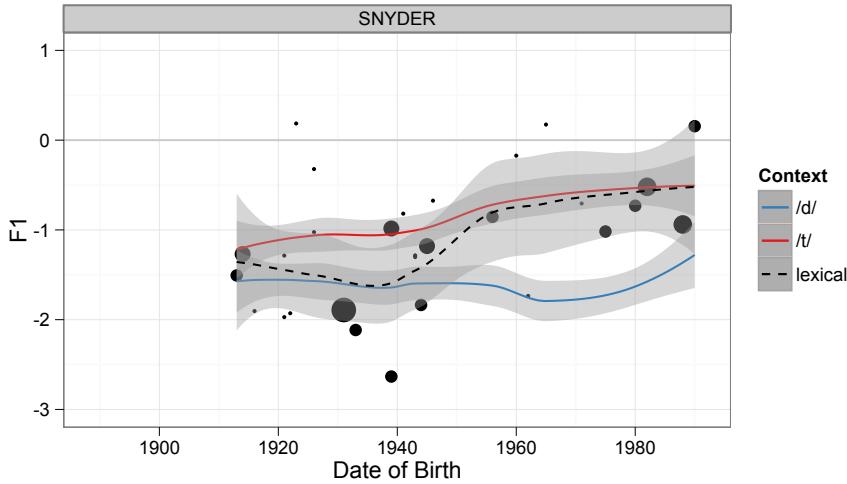


Figure 28: Exceptional raising, compared to raising before /t/ and /d/.

pound, or secondary stress on the second syllable. Figure 29 displays this comparison. The data in the *Friday* panel represents the mean height of *Friday* per-speaker, and the average height of /ayt/ and /ayd/ for those speakers who contributed *Friday* tokens. Note that speakers are not necessarily represented in both panels. In fact, of the 64 speakers who contributed *Friday* tokens, and of the 31 speakers who contributed SNYDER class tokens, only 6 contributed tokens to both.

Based on Figures 20 and 21, /ay/-raising began at some point around the turn of the century, and was well under way by 1910 at the latest. There was approximately a 30 year gap between the opacity of /ay/-raising, and the beginning of lexical diffusion of raised /ay/. More data is necessary to be certain, however. Part of this proposal is to collect more data, specifically on the lexical diffusion of /ay/ raising, targeting speakers born between 1940 and 1960.

Additional Case Studies

I will not be discussing additional case studies in as much detail here as I did /ay/-raising; however, there are a few more conditioned sound changes in Philadelphia which will be part of the dissertation.

I have already done some work on the raising of /ey/, which was also identified as a new and vigorous change in the 1970s. The results of that case study have shown that the

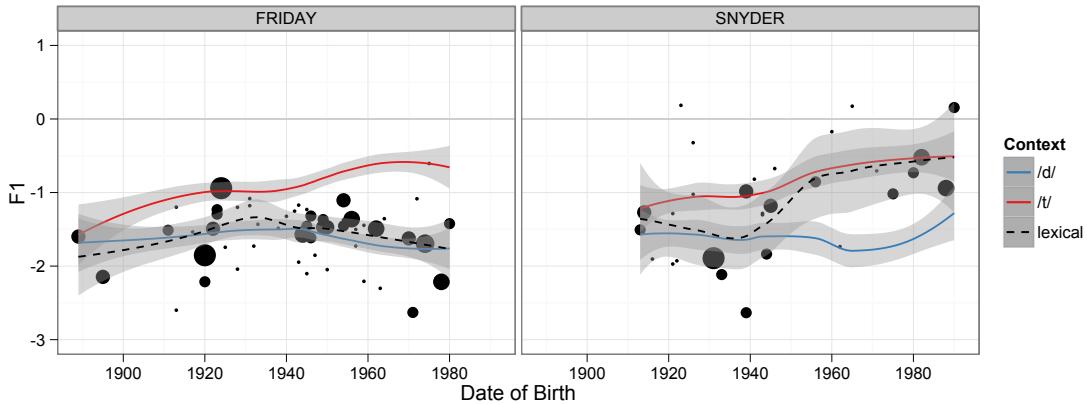


Figure 29: Contrasting SNYDER class with *Friday*.

conditioning of the change can be described as /ey/ followed by a consonant within the same word, since /ey/ does not raise word finally, nor pre-hiatus. The status of /l/ with regards to this conditioning is particularly interesting, as it seems to phonetically favor a slightly raised /ey/, but does *not* condition the change. I also found no interaction with /ey/-raising and inflectional phonology, i.e. *day*=*days*=*daze*. The data in the PNC was too sparse to investigate the effect of derivational morphology, or compounding, which would be another avenue of investigation to pursue with additional data collection.

Another conditioned change which I have not yet done any analysis of is the fronting of /uw/, which has two conditioning factors. First is whether or not the preceding consonant is apical, and second is following /l/. The combination of two conditioning factors could provide some crucial data for distinguishing between phonetic and phonological conditioning. If both preceding apicals and following /l/ are phonetic conditioning factors, then we would expect their effects to combine in an additive way. However, if the conditioning is largely phonological, we might expect to see one cancel out the other. A deep investigation into phonological structure seems to be in order for this analysis, because my own intuition is that fronting is blocked in words like *tooling*, but not in *tulip*, the distinction being the status of the /l/ as an onset or coda at different phonological strata (Bermúdez-Otero, 2011; Turton, 2011).

Further Directions

There are two issues which I have not yet addressed, which will have to be firmly hammered out in the dissertation. First, I have been saying that there is a phonological distinction between /ay[+voice]/ and /ay[–voice]/, but I have not yet attempted to define what that distinction is. An additional complication here is that the phonological distinction would necessarily have a very small phonetic difference near the beginning of the change. The easiest approach would be to use a diacritic feature, but ideally there would not be just one phonological feature which distinguishes between two allophones of /ay/, and nothing else. I should, then, attempt to find any concomitant phonetic effects of /ay/-raising in other vowels which might share the relevant feature with /ay[–voice]/.

The second remaining issue is how a phonological process which creates a new allophone of /ay/ entered into the dialect apparently overnight. Unlike the phonetic change, the introduction of a phonological process *is* a discrete process, and we would expect it to have the same dynamic as other discrete changes, as discussed in §2. As I illustrated in §2, /ay/-raising does not look like a change that progresses via competition between two categories with very different phonetics. However, it may be possible that in the early stages of the change, there is competition between two categories with very similar phonetics. This phonological change would have to have finished very early in order for most of the change to look like just the continuous movement of one category through phonetic space, but it might be possible for a phonological change like this to go from beginning to completion in 20 years. The lexical diffusion of the SNYDER class, for example, took place along this kind of time span, as did the shift of Montreal /r/ from [r] to [ɹ] (Sankoff and Blondeau, 2007). Determining whether or not there was categorical variation between two phonetically similar categories would require a manual analysis of speakers. I would begin with speakers born between 1888 and 1910 (26 speakers), and then move into later dates of birth as necessary.

It might also be worth while to contrast patterns in conditioned changes to unconditioned changes. For example, I have been arguing that a phonological reanalysis of /ay[–voice]/ occurred at the onset of /ay/-raising. What, then, is the status of a change like /aw/-raising? /aw/-raising has no phonological conditioning, meaning that there is no allophone which does not undergo the change. Either /aw/ was phonologically reanalyzed before it began to raise, like /ay[–voice]/ was, or unconditioned changes have different sequencing of phonetic shifts and phonological reanalysis. The sequencing of phonetic and phonological change is especially important to understand when it comes to a phenomenon like the Northern Cities Shift, where the phonological system has clearly been reorganized (Labov et al., 2006; Preston, 2008).

3.2 Continuing

Once the early influence of phonology on phonetic change has been established, the next question to ask is whether phonology has only an early influence. For instance, the phonological influence might be, to follow the metaphor utilized by Janda and Joseph (2003), a “Big Bang,” or a sudden event that sets the direction of the phonetic change, which is then driven along by its own inertia, not the persistent influence of phonology.

My focus here is specifically on the core, or original phonetic change. It is clear that new, or additional changes can be overlaid on top of the original change. I have already discussed one such case in §3.1 with the lexical diffusion of raised-/ay/ to new words. Phonological conditioning factors may also be altered during the course of a change, either by expanding or narrowing (Janda and Joseph, 2003; Bermúdez-Otero and Trousdale, 2011).

However, what I am focusing on here is whether or not phonology has a continuing relationship with phonetic change at all time points, rather than just during punctuated events like the actuation, expansion, or narrowing of a change. To support the argument that there is a continuing phonological effect throughout the course of a phonetic change, I'll focus on two pieces of evidence. First are some preliminary results that /ay[+voice]/ is decreasing in duration as /ay[–voice]/ is raising in height, suggesting a cue-trading pattern. Secondly, I'll discuss patterns in Parallel Shifts which suggest that the units of phonetic change can be smaller than a phoneme, targeting a feature, or subset of features, which in turn affects multiple phonemes.

/ay[+voice]/ Shortening

The first pieces of evidence I have of a continuing, rather than punctual, influence of phonology on phonetic change are preliminary results that /ay[+voice]/ is shortening while /ay[–voice]/ is raising. There is a hint of decreasing duration for /ayd/ in Figure 22 above. There is also a large main effect of prosodic context, which must be controlled for in such an investigation. For these preliminary results, I am only looking at /ayd/ sequences, and will expand the analysis to all /ay[+voice]/ sequences in the dissertation. An analysis of duration must be carefully done, not only because of the prosodic main effect, but also segmental main effects (House and Fairbanks, 1953), meaning the analysis may need to be done segment by segment. And again, the duration data available to me from the PNC is censored at 50ms, making it necessary for me to re-collect the duration data from all /ay/’s.

For /ayd/, I defined 5 contexts that could have an effect on duration.

1. C#

/ay/, followed by /d/, followed by pause.

2. internal flap
/ay/, followed by /d/, followed by an unstressed vowel within the same word.
3. boundary flap
/ay/, followed by /d/, followed by a word boundary, followed by a vowel.
4. vcc#
/ay/, followed by /d/, followed by another consonant within the same word.
5. vc#c
/ay/, followed by /d/, followed by a word boundary, followed by a consonant.

I fit two mixed effects models predicting log of duration by date of birth. One model had both a main effect of context and an interaction with date of birth, and the other had only a main effect. The simpler model was preferred by the AIC ($4398 - 4404 = -6$) and BIC ($4451 - 4481 = -30$), meaning it is more likely that there is no interaction between date of birth and context. The model predicts a $2^{-0.03 \times \text{decade}}$ rate of decrease in duration, meaning that over the century of data we've collected, the duration of /ayd/ has decreased by about 20%. Figure 30 displays the vectors of change across duration and F1.

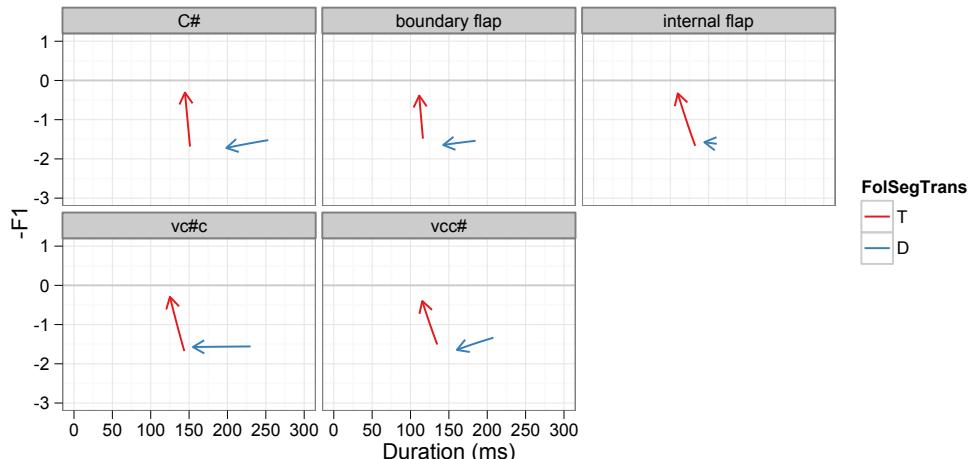


Figure 30: /ay/ duration and height.

This pattern looks like diachronic cue trading, similar to what has been observed synchronically (e.g. vowel stress (Howell, 1993)). Some phonological contrast which used to be cued primarily with duration began to be cued by vowel quality, which then led to the reduction of the duration cue. It might be asked which phonological distinction

is being cued: the voicing of the following segment, or the allophone of the vowel. If the pre-voiceless shortening effect in Philadelphia were phonological, this would allow for a phonological distinction between /ay[+voice]/ and /ay[–voice]/ to have existed much earlier than 1900, perhaps simplifying the problem of how the specification of /ay/-raising spread so rapidly throughout the speech community.

Parallel Shifts

A class of phonetic shifts which are possibly the best examples of phonological involvement in phonetic change are Parallel Shifts. They are not classified in the taxonomy of Chain Shifts in Labov (1994), perhaps because they are not chained sequences of events exemplified by gap creation and filling, but rather as the parallel movement of a natural class. The most prominent contemporary parallel shift is the Canadian Shift, as described by Boberg (2005) in Montreal, and Durian (2009) in Columbus, OH. The original descriptions of the shift (Clarke et al., 1995; Labov et al., 2006) stated that first, /æ/ retracted, followed by /ɛ/ lowering into the position vacated by /æ/, then /ɪ/ lowering into the position vacated by /ɛ/. Boberg (2005) and Durian (2009), on the other hand, describe a parallel retraction of /æ/, /ɛ/, and /ɪ/. Figures 31 and 32 illustrate.

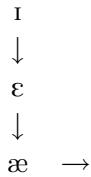


Figure 31: Canadian Chain Shift

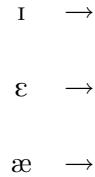


Figure 32: Canadian Parallel Shift

Neither Boberg (2005) nor Durian (2009) attempt to account for the parallel version of the shift in terms of phonology, although Durian (2009) does promote a view of it as “phonetic analogy.” However, it seems clear that the set of vowels undergoing the shift is readily defined as a phonological natural class defined by [–back] and [–peripheral], and it exactly along these phonetic dimensions that they are shifting in parallel. Parallel shifts of this variety speak strongly for a continuing influence of phonology throughout a phonetic change because the definition of the set of segments undergoing the change is a phonological natural class. That is, the phonetic change is directly related to the phonological features [–back] and [–peripheral].

None of the active phonetic changes in Philadelphia have been described in terms of parallel shifts before, but it may be possible to do so. In particular, the fronting of the long, upgliding vowels /uw, ow, aw/, and the parallel raising of /ey, ay/ are possible candidate cases for reanalysis as parallel shifts, and I will provide preliminary results to

support that here. Part of the dissertation work would be to evaluate the reasonableness of such an analysis, and how it interacts with other factors, such as the social correction of /aw/.

Figures 33, 34, and 35 constitute the primary piece of evidence for the parallelism of these changes. Each point represents a speaker, and the relationship between their /uw/, /ow/ and /aw/ frontness (frontness and height in the case of /aw/). The data used to create these graphs excludes specific conditioning environments for each vowel, to focus on the core pattern.^{7,8} The points are colored according to sex because that has been found to be an important dimension of social stratification for /aw/. Kendall's τ , a non-parametric measure of correlation, is also displayed for men and women, along with a p-value for a two-tailed test of the null hypothesis that $\tau = 0$.⁹ The weakest relationship appears to be between /aw/ and /uw/, which also differ by two degrees of height. The correlation is not significant for men, and is the weakest of the correlations from women. The relationships which only differ by one degree are much stronger, especially /uw/ and /ow/. An interesting outcome from splitting the data according to sex is that the correlations are much higher for women than for men for all three comparisons. If I find that these patterns are strong and convincing, my argument would frame this as a phonetic change effecting the feature which defines these three vowels as a natural class.

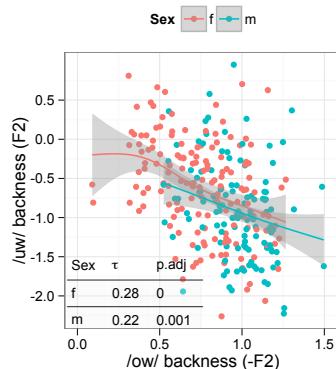


Figure 33: /ow/ ~ /uw/

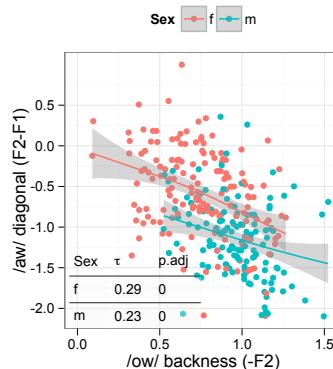


Figure 34: /ow/ ~ /aw/

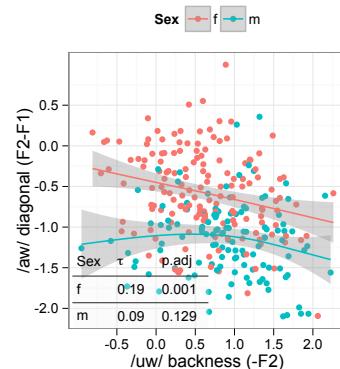


Figure 35: /uw/ ~ /aw/

⁷Specifically, /uw/ following apicals, /ow/ in final position, /aw/ before nasals, and all three before /l/ were excluded.

⁸Additionally, the dimension along which /uw/ and /ow/ are plotted is simply F2, while /aw/ has been plotted along F2-F1, which we have found to be a good measure of movement along the front diagonal.

⁹These p-values have been adjusted by the Holm-Bonferroni method, whereby the smallest p-value is multiplied by n , where n = the number of tests. The second smallest p-value is then multiplied by $n - 1$, the third smallest by $n - 2$, etc. p-values of 0 indicate that the $p < 0.001$, and was too small to be rounded up to 0.001.

The second candidate for a parallel shift is the raising of /ey/ and /ay/. These are first and second most vigorous shifts in the PNC, they are both front up-gliding vowels, and there are a number of similarities in their conditioning. /ey/-raising is conditioned by any following consonant, while /ay/-raising is more specifically conditioned by a following voiceless consonant. In both cases, the conditioning environment is restricted to be within the word level stratum. Furthermore, both changes begin around the same time period. A sketch of what the phonological process involved here is given in (3).

$$(3) \quad V_y \rightarrow V_y^+ / \underset{(-\text{voice})}{C} l_\omega$$

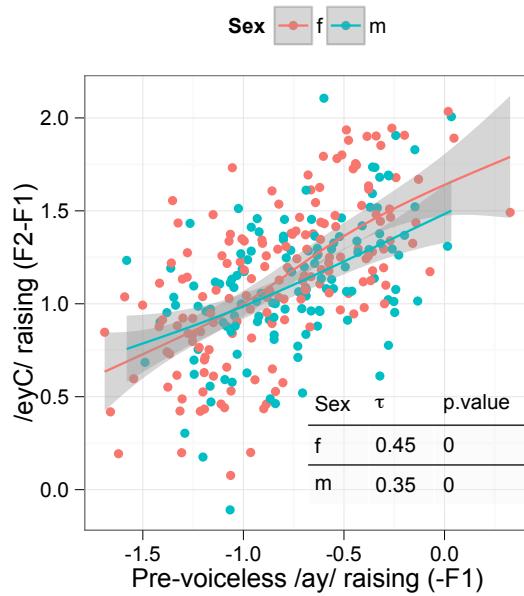


Figure 36: /ay[–voice]/ ~ /eyC/

Figure 36 plots the relationship between /eyC/ and /ay[–voice]/ across individuals. Clearly, the relationship is very tight. The reason that I am arguing this is a parallel shift, rather than a chain shift, is because /ay[–voice]/ is phonetically central: [ʌi], which means it is not moving into a phonetic space vacated by /eyC/: [eɪ]. Furthermore, this would be a strange example of an allophonic chain shift, which is not attested to occur, and at least the investigation of one case study in Labov (2010) found no evidence for one. Rather, this could be argued to be a parallel phonetic shift affecting an allophone of /ay/ and an allophone of /ey/ which are defined by the same phonological feature.

Conclusion

Both the apparent diachronic cue trading for allophones of /ay/ and the existence parallel phonetic shifts suggest that phonetic change is mediated by phonological representations.

4 The Mechanics of Phonetic Change

To briefly recap, the facts of phonetic change as they have been laid out here is that it clearly progresses as a continuous shift through the phonetic space, but is also heavily mediated by categorical phonological representation early, and throughout the change. These facts raise the theoretical question of *what* is changing. A detailed answer to this question depends heavily on one's hypothesis about the relationship between phonetic and phonological representation. A general answer, however, is that the phonetic realization of a phonological category changes, while the category itself remains relatively stable. This fact immediately brings the nature of the phonology-phonetics interface to bear on the process of phonetic change, and vice versa.

The facts of phonetic change demand language specific phonetics of some sort. The most basic sketch of a system of language specific phonetics is Figure 37, which is adapted from Keating (1990). The phonological system maps an underlying phonological form to a surface phonological form. This surface phonological form is then mapped to a phonetic representation.

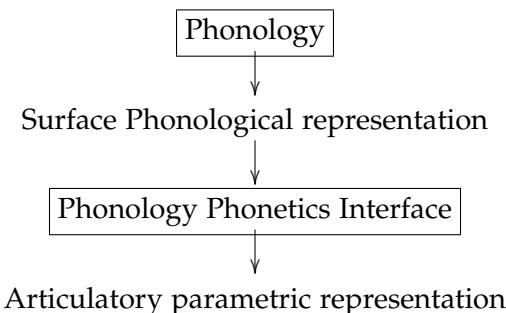


Figure 37: Model of Language Specific Phonetics

4.1 Objects of Implementation

This very basic sketch leaves a great many important questions unanswered. First, what is the size of the surface phonological object that gets implemented? It could be just a feature, like [-back], and all objects which share this feature also share a similar phonetic

implementation. It could be an atomic, or wholistic representation of a segment, which are typically represented with IPA symbols, like /æ/. Finally, it could be a sequence of segments implemented together, allowing for some segments to have unique phonetic realizations in the context of others.

Cohn (1993) represents an argument for the first possibility: feature-to-phonetics mapping. She compares coarticulatory nasalization and nasal assimilation. Segments are represented as either [\pm nasal], or unspecified for nasality, depending on the phonological circumstances. Nasal airflow remains at a minimum throughout segments specified [–nasal], and transitions sharply to high airflow if followed by a segment specified [+nasal]. Nasal airflow can transition gradually across segments which are unspecified for nasality. All segments which share a [–nasal] or [+nasal] specification share a rate of nasal airflow in their phonetic implementation. The cases of parallel shifts like the Canadian Shift are also good evidence for feature-to-phonetics mapping, where the phonetic correlate of [–back] is changing over time, affecting the [–back] set of vowels.

There are few examples of explicit arguments for the second possibility: segment-to-phonetics mapping. However, most discussion within sociolinguistics and dialectology, especially those that frame their analyses within Exemplar Theory, implicitly describe their observations in terms like these. Even, Durian (2009) describes the Canadian Shift in Columbus, Ohio as parallel /æ/-retraction, /ɛ/-retraction, and /ɪ/-retraction. I have argued that this shift should be reconsidered as a parallel, feature-to-phonetics change, but in fact, there are a great many phonetic changes in the sociolinguistic literature where only a singleton segment is affected. The influence of Exemplar Theory on sociophonetics has especially increased the tendency to describe segment specific changes, because segment-to-phonetics mapping can be achieved by associating a vowel category label with a number of phonetic memory traces of that category.

The phonetic implementation of a sequence of segments is, perhaps, the least theoretically satisfying option. However, Boersma and Hamann (2008) proposed something like this when discussing pre-voiceless shortening using bidirectional cue constraints. A bidirectional cue constraint can be described as a constraint on a tuple of phonetic realization, and phonological representation: *⟨phonetic realization, phonological representation⟩. It can be read as “this phonetic realization for this phonological representation is marked.” It is a bidirectional constraint, because it can be used both in production, (phonological representation input, phonetic realization output) and perception (phonetic realization input, phonological representation output). For pre-voiceless shortening, Boersma and Hamann (2008) proposed the cue constraint in (4). It can be read as “a long duration for a vowel is marked when followed by a phonologically voiceless segment.”

- (4) *⟨long duration, V⟩⟨ __, –voice⟩

To some extent, this contextual phonetic implementation constraint is difficult to dis-

tinguish from a phonological process which adds a diacritic to the vowels preceding voiceless consonants.

(5) Phonology: $V \rightarrow V^+ / _ -\text{voice}$

(6) Phonetics: *⟨long duration, V^+ ⟩

However, it remains a conceptual possibility.

Being explicit about the size of phonological representations which serve as input to the phonology-phonetics interface is important if I am going to treat phonetic change as changing phonetic implementation. The units of phonetic implementation should be, under such an approach, the same as the units of phonetic change. If segments can undergo phonetic shifts as atomic units (i.e. as a singleton set, leaving behind other members of its natural class), then that would be good evidence for segments, as units of phonetic implementation. If natural classes can undergo shifts together, that would be good evidence for features as units of phonetic implementation. Framed this way, phonetic change can serve as direct evidence on the nature of the phonology-phonetics interface.

5 Timeline

Summer 2012

- Collect extra necessary data from the PNC, including
 - More complete duration information.
 - More measurement points, ideally glide targets.
- Establish the best statistical practices for dealing with semi-longitudinal data.
- Begin making introductions for additional interviews.

September-October 2012

- Finalize interview modules and formal methods, conduct first interviews.
- Literature review on phonetic change and phonological change.
- Revise and refine case studies of conditioned sound changes.
 - /ay[–voice]/
 - /eyC/

- /uwl/
- Chapter 1: Introduction

November-December 2012

- Complete and transcribe interviews.
- Literature review on phonology-phonetics interface.
- Chapter 2: Models of Phonetic Change
- Chapter 3: Data and Analysis Methods
 - The Philadelphia Neighborhood Corpus
 - Additional interviews

January-February 2013

- Chapter 4: Conditioned Phonetic Changes
 - ay0
 - eyC
 - owl
- Chapter 5: Parallel Phonetic Changes

March-April 2013

- Chapter 6: Consequences for Phonological and Phonetic Theory
- Chapter 7: Conclusions

May 2013

- Final revisions.
- Defend & submit dissertation.

References

- Baker, Adam, Diana Archangeli, and Jeff Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23:347—374.
- Bermúdez-Otero, Ricardo. 2007. Diachronic Phonology. In *The Cambridge handbook of phonology*, ed. Paul de Lacy, 497–517. Cambridge: Cambridge University Press.
- Bermúdez-Otero, Ricardo. 2011. Cyclicity. In *The Blackwell Companion to Phonology*, ed. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice. Wiley-Blackwell.
- Bermúdez-Otero, Ricardo, and Graeme Trousdale. 2011. Pathways of phonological change. lingBuzz/001235.
- Blevins, Juliette. 2004. *Evolutionary phonology: the emergence of sound patterns*. Cambridge University Press.
- Boberg, Charles. 2005. The Canadian shift in Montreal. *Language Variation and Change* 17:133–154.
- Boersma, Paul, and Silke Hamann. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25:217–270.
- Braver, Aaron. 2011. Incomplete Neutralization in American English Flapping: A Production Study. In *University of Pennsylvania Working Papers in Linguistics*, ed. Lauren Friedman, volume 17.1, 31–40.
- Brody, Stefanie. 2009. The status of planets: development and transmission of Philadelphia short-a. Paper presented at PLC 33.
- Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14:261–290.
- Cedergren, Henrietta J., and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50:333–355.
- Chambers, J. K. 1973. Canadian raising. *Canadian Journal of Linguistics* 18:131–135.
- Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22:129–159.
- Clarke, Sandra, Ford Elms, and Amani Youssef. 1995. The Third Dialect of English: Some Canadian evidence. *Language Variation and Change* 7:209–228.

- Cohn, Abigail. 1993. Nasalisation in English: Phonology or Phonetics. *Phonology* 10:43–81.
- Disner, Sandra. 1978. *Vowels in Germanic Languages*, volume 40 of *Working Papers in Phonetics*. UCLA Department of Linguistics.
- Durian, David. 2009. Purely a chain shift?: An exploration of “Canadian Shift” in the US Midland. Paper presented at NNAV 38.
- Evanini, Keelan. 2009. The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania. Doctoral Dissertation, University of Pennsylvania.
- Fruehwald, Josef. 2007. The spread of raising: Opacity, lexicalization, and diffusion. *CUREJ: College Undergraduate Research Electronic Journal*.
- Fruehwald, Josef, Jonathan Gress Wright, and Joel Wallenberg. 2009. Phonological rule change: The constant rate effect. Paper presented at NELS 40.
- Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. New York: Oxford University Press.
- House, Arthur S., and Grant Fairbanks. 1953. The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels. *The Journal of the Acoustical Society of America* 25:105–113.
- Howell, Peter. 1993. Cue trading in the production and perception of vowel stress. *Journal of the Acoustical Society of America* 94:2063–2073.
- Huwaldt, Joseph A. 2011. Plot digitizer v. 2.5.1. Software.
- Idsardi, William J. 2007. Canadian Raising, Opacity and Rephonemicization. *Canadian Journal of Linguistics* 51(2/3):119–126.
- Janda, Richard D., and Brian D. Joseph. 2003. Reconsidering the Canons of Sound-Change: Towards a “Big Bang” Theory. In *Historical Linguistics 2001. Selected Papapers from the 15th International Conference on Historical Linguistics, Melbourne*, ed. Barry Blake and Kate Burridge, 205–219. John Benjamins.
- Joos, Martin. 1942. A phonological dilemma in Canadian English. *Language* 18:141–144.
- Keating, Patricia. 1990. Phonetic representations in a generative grammar. *Journal of Phonetics* 18:321–334.
- King, Robert D. 1973. Rule insertion. *Language* 49:551–578.

- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.
- Labov, William. 1969. Contraction, Deletion, and Inherent Variability of the English Copula. *Language* 45:715–762.
- Labov, William. 1981. Resolving the Neogrammarian Controversy. *Language* 57:267–308.
- Labov, William. 1994. *Principles of linguistic change. Volume 1: Internal Factors*. Blackwell.
- 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*. Blackwell.
- Labov, William, Sherry Ash, and Charles Boberg. 2006. *The Atlas of North American English. Phonetics, Phonology and Sound Change..* Berlin: Mouton de Gruyter.
- Labov, William, Paul Cohen, Clarence Robins, and John Lewis. 1968. *The Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City: Phonological and grammatical analysis*.
- Labov, William, and Ingrid Rosenfelder. 2011. The Philadelphia Neighborhood Corpus.
- Labov, William, Malcah Yaeger, and Richard Steiner. 1972. *A Quantitative Study of Sound Change in Progress*. The U.S. National Survey.
- Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. New York: Oxford University Press.
- Moreton, Elliott, and Erik R. Thomas. 2004. Origins of Canadian Raising in voiceless-coda effects: a case study in phonologization. In *Papers in Laboratory Phonology 9*, ed. J. Cole and J. I. Hualde. Cambridge University Press.
- Moreton, Elliott. 2004. Realization of the english postvocalic [± voice] contrast in f1 and f2. *Journal of Phonetics* 32.
- Ohala, John. 1981. The listener as a source of language change. In *Papers from the parasession on language and behavior*, ed. C. S. Masek, R. A. Hendrick, and M. F. Miller. Chicago Linguistics Society.
- Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency and the emergence of linguistic structure*, ed. Joan Bybee and Paul Hopper, volume 45 of *Typological Studies in Language*, 137–157. John Benjamins.

- Pitt, M. A., L. Dilley, Keith Johnson, S. Kiesling, W. Raymond, Elizabeth Hume, and E. Fosler-Lussier. 2007. Buckeye corpus of conversational speech (2nd release).
- Preston, Dennis R. 2008. What is a Speech Community Anyhow. Talk presented at NWA37.
- Rosenfelder, Ingrid. 2005. Canadian Raising in Victoria, B.C.: An Acoustic Analysis. Master's thesis, University of Heidelberg.
- Sankoff, David, and Gillian Sankoff. 1973. Sample survey methods and computer assisted analysis in the study of grammatical variation. In *Canadian languages in their social context*, ed. R. Darnell. Linguistic Research.
- Sankoff, Gillian, and Hélène Blondeau. 2007. Language Change across the Lifespan: /r/ in Montral French. *Language* 83:560–588.
- Tauberer, Joshua. 2010. Learning [voice]. Doctoral Dissertation, University of Pennsylvania.
- Tucker, R. Whitney. 1944. Notes on the Philadelphia Dialect. *American Speech* 19:37–42.
- Turton, Danielle. 2011. The Darkening of English /l/: a Stochastic Stratal OT Analysis. Paper presented at MFM 19.
- Wedel, Andrew. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23:247–274.
- 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.
- Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12:231–250.
- Yang, Charles. 2002. *Knowledge and learning in natural language*. New York: Oxford University Press.
- Yuan, Jiahong, and Mark Liberman. 2008. Speaker identification on the SCOTUS corpus. In *Proceedings of Acoustics*.