

THE PHONOLOGICAL INFLUENCE ON PHONETIC CHANGE

Josef Fruehwald

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2013

William Labov, Professor of Linguistics, Supervisor of Dissertation

Eugene Buckley, Associate Professor of Linguistics, Graduate Group Chairperson

Dissertation Committee:

Charles Yang, Associate Professor of Linguistics

Ricardo Bermúdez-Otero, Senior Lecturer in Linguistics and English Language, University of Manchester

THE PHONOLOGICAL INFLUENCE ON PHONETIC CHANGE

COPYRIGHT

2013

Josef Fruehwald

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0
Unported License.

Acknowledgements

I've lived in Philadelphia for 28 years, been at Penn for 10, and hanging out in and around the Linguistics Department for 9. The fact that this chapter of my life is now approaching its conclusion is bittersweet. The fact that I am happy to go but sad to say goodbye is a testament itself to how lucky I have been, but it only begins to scratch the surface. Over the past decade, I have not only discovered my personal heroes, but also had the supreme good fortune to have them as teachers, fellow students, mentors, and peers. Viktor Frankl said that humanity has a "will to meaning," that above all, people desire to have a direction and passion in life. The UPenn Linguistics department and its members have given me no less than that.

Among people I should thank individually, my wife, Becky Mead, tops the list. I could thank her for her patience when I fell behind in my household responsibilities, or for her support and planning acumen as we begin planning the next stage of our life, but above all, I'd like to thank her for keeping it a secret that she'd bought plane tickets for two to Japan when I was too nervous that I'd fail my proposal to commit to going. She has an unflagging confidence in me, even when I am plagued by insecurities, and that has done more to get me through this process than I can express.

William Labov and Gillian Sankoff have been the very definition of mentors to me. I learned from them that I had the ability to contribute new bits of knowledge to the field of linguistics, starting with Intro to Sociolinguistics, which Gillian taught in my sophomore year at Penn. I've also learned from their doing how to appreciate the intelligence and ability of non-experts, and of experts I disagree with ("Linguists are smart," Bill says), both characteristics which are not naturally the product of a graduate education. They both took my proposal for my senior thesis so seriously and enthusiastically, when at the time I was still a relative neophyte. I owe my present day successes to their respect and support back then, and I look forward to encouraging mentees of my own to pay it forward. All in all, there could not have been a better place for a Philly boy with just a nascent social consciousness to land than in the Penn Linguistics Lab.

I should also thank the members of my committee for their support and input. This research project began as a term paper I wrote for a seminar Gene Buckley taught. At that point in time, it was *rough*, but has gotten better with the help of Gene's keen eye and insightful questions. Charles Yang's level headed approach to integrating probabilities and statistics into a theory of language has been an inspiration in my own work, and I can link thinking on many matters covered in Chapter 6 directly to meetings, classes, and conversations with him. One of the greatest compliments you can give an academic is to tell them that you read what they wrote, and that you grasped what you meant to say. On this front, Ricardo Bermúdez-Otero is one of the greatest complimenters there is, and both his nearly encyclopedic knowledge of the literature, and his

close reading of my own writing has pushed this dissertation, and my own thinking, further than it would have been otherwise.

I'd like to single out Joel Wallenberg for thanks. Originally, Joel was my TA. When I started grad school he was a peer, and we rapidly became close friends. I have fond memories of wandering down 45th street in West Philly having long and rambling conversations with Joel on every topic imaginable. The nicest thing about conversations with him is that I can propose a half formed idea, and he can finish it, and vice versa, mutually reassuring ourselves that we're not crazy people, and might actually be on to something. At least, so we tell each other, and if we're totally wrong, I couldn't think of a better person to be totally wrong with.

Since the Linguistics Lab has been my academic home for about 6 or 7 years now, I'd like to thank the people who have made it such a wonderful social and intellectual environment, especially Sue Sheehan. I don't know what we'd do without Sue. She takes on tasks that I know would fill my own heart with dread. The Philadelphia Neighborhood Corpus, and thus my research, has been made possible largely though Sue's hard work and dedication. And who else can in one instant be keeping diligent track of the recordings in the lab's tape closet, and in the next chase down the university president's personal assistant and demand they set up the generators for the garden party elsewhere because they're blowing fumes through our windows and disrupting our delicate intellectual process? Also, special thanks go to my officemates, Laurel MacKenzie, Ingrid Rosenfelder, Meredith Tamminga, and Hilary Prichard, those productive people whose flow I repeatedly interrupt with topical and non-topical interjections.

Of course, I have made a habit of interrupting productive people not only at the Linguistics Lab, but also at the Institute for Research in Cognitive Science. So thank you to Andy Gersick, Elika Bergelson, Constantine Lignos, Kyle Gorman, Aaron Ecay and Chris Ahern for putting up with me. John Trueswell and Christine Massey have been extremely supportive in their roles as IGERT PIs. I'd also like to thank the regular attendees of the Common Ground seminar for both broadening my horizons, and for their thoughtful feedback on my own work, especially Stephen Isard.

There are too many people in the Linguistics department to name and thank individually, but I'd like to start with my undergrad TAs, including Damien Hall, Aaron Dinkin, Łukasz Abramowicz, Michael Friesner, Keelan Evanini and Joshua Tauberer. My extended conversations with Tony Kroch and Dave Embick have been invaluable for the development of my thinking about language, and for bringing together data and theory. I should also thank Amy Forsyth for her strongly worded e-mails, keeping me on track with the program. One of the most valuable aspects of my graduate school experience is the degree to which we as grad students learned from and supported each other. I owe everyone I was ever in class with, TAed with, was on a PLC committee with, or really participated in any kind of project with, for contributing to our collective knowledge and morale. On my first day, the graduate school dean told us at an orientation that grad school is ultimately a solitary experience. I think he was dead wrong, and I couldn't be more glad.

I'd also like to briefly thank Hadley Wickham and Yihui Xie for the tools they've built that have made it possible for me to write this dissertation. I've never met them in person, but I do follow them on Twitter.

Last, but certainly not least, I'd like to thank my parents, Franz and Marian Fruehwald, and my siblings, Paul, Kate and Rebecca Fruehwald. One of the more surprising aspects of entering adulthood is realizing that your family has played an even bigger role in shaping you into the person you are than you had previously realized. I would not be the person I am today if not for

their love and support.

These thanks and acknowledgements are incomplete. Given the time and space available, there is no way for me to sufficiently thank everyone who has played a significant part in me completing my dissertation. For those who haven't been mentioned by name, and for those whose love and support I've understated, you have my deepest and sincere thanks and apologies.

ABSTRACT

THE PHONOLOGICAL INFLUENCE ON PHONETIC CHANGE

Josef Fruehwald

William Labov

This dissertation addresses the broad question about how phonology and phonetics are interrelated, specifically how phonetic language changes, which gradually alter the phonetics of speech sounds, affect the phonological system of the language, and vice versa. Some questions I address are:

- (i) What aspects of speakers' knowledge of their language are changing during a phonetic change?
- (ii) What is the relative timing of a phonetic change and phonological reanalysis?
- (iii) Can a modular feed-forward model of phonology and phonetics account of the observed patterns of phonetic change?
- (iv) What are the consequences of my results for theories of phonology, phonetics, and language acquisition?
- (v) What unique insight into the answers to these questions can the study of language change in progress give us over other methodologies?

To address these questions, I drew data from the Philadelphia Neighborhood Corpus [PNC] (Labov and Rosenfelder, 2011), a collection of sociolinguistic interviews carried out between 1973 and 2013. Using the PNC data, I utilized a number of different statistical modeling techniques to evaluate models of phonetic change and phonologization, including standard mixed effects regression modeling in R (Bates, 2006), and hierarchical Bayesian modeling via Hamiltonian Monte Carlo in Stan (Stan Development Team, 2012).

My results are challenging to the conventional wisdom that phonologization is a late-stage reanalysis of phonetic coarticulatory and perceptual effects (e.g. Ohala, 1981). Rather, it appears

that phonologization occurs simultaneously with the onset of phonetic changes. I arrive at this conclusion by examining the rate of change of contextual vowel variants, and by investigating mismatches between which variants are expected to change on phonetic grounds versus phonological grounds. In my analysis, not only *can* a modular feed-forward model of phonology and phonetics account for observed patterns of phonetic change, but must be appealed to in some cases.

These results revise some the facts to be explained by diachronic phonology, and I suggest the question to be pursued ought to be how phonological innovations happen when there are relatively small phonetic precursors.

Contents

Acknowledgements	iii
Abstract	vi
Contents	viii
List of Tables	x
List of Figures	xii
1 Introduction	1
2 What is Phonetic Change?	6
2.1 Sound Change and Grammar	6
2.1.1 Phonemic Incidence	9
2.1.2 Systems of Phonological Contrast	12
2.1.3 Presence or Absence of Phonological Processes	14
2.1.4 Targets of Phonetic Implementation	16
2.1.5 Gestural Phasing and Interpolation	18
2.1.6 Sound Change and Grammar Summary	19
2.2 The Phonology-Phonetics Interface	20
2.2.1 Modular and Feedforward	20
2.2.2 The Architecture	21
2.2.3 Sociolinguistic Variation	28
2.3 Phonetic Change	30
2.3.1 Phonetic Change is Continuous	30
2.4 Conclusion	42
3 The Philadelphia Neighborhood Corpus Data	44
3.1 The Philadelphia Neighborhood Corpus	44
3.1.1 Forced Alignment and Vowel Extraction (FAVE)	45
3.2 Enrichment of Contextual Information	47
3.3 Total Data Count	48
3.4 Normalization	48
3.5 Choice of Time Dimension	49

4 The Rate of Phonetic Change	54
4.1 Phonetic Coarticulation vs Phonological Differentiation	57
4.1.1 Phonological vs Phonetic Processes in Sound Change	61
4.2 The Rate of Language Change	65
4.3 The Model and the Data	69
4.3.1 The Model	69
4.3.2 Implementing the model	77
4.3.3 I have just described a generative model	79
4.4 Case Studies	80
4.4.1 /aw/	81
4.4.2 /ow/	88
4.4.3 /uw/	97
4.5 Summary of /Vw/ results	105
4.5.1 Connection to Broader Theory	107
4.6 Conclusion	108
5 Phonologically conditioned divergence and convergence	110
5.1 Phonologically divergent behavior within categories	110
5.1.1 /ay/ Raising and Opacity	110
5.1.2 /ey/ Raising	131
5.1.3 /ay/ and /ey/ summary	140
5.2 Natural Class Patterns	141
5.2.1 Back vowel fronting in Philadelphia	145
5.2.2 Long-ingliding vowel lowering in Philadelphia	154
5.2.3 Searching for more parallel shifts	157
5.2.4 Parallel Shifts are Changing Phonetic Implementations of Phonological Features, but there are Complications	161
5.3 Conclusions	164
6 Against Gradual Phonologization	165
6.1 Conventional Wisdom Regarding Sound Change	165
6.1.1 This conventional wisdom across research programs	168
6.1.2 The challenge posed by my results	172
6.2 Big Bang	179
6.2.1 Plausibility	183
6.2.2 Big Bang Summary	189
6.3 Similarity to syntactic change	189
6.4 Additional Challenges, and Directions for Future Research	191
7 Conclusions	193
Bibliography	196

List of Tables

3.1	Model comparisons for using different time dimensions to predict pre-voiceless /ay/ height	52
3.2	Model comparisons for using different time dimensions to predict /aw/ raising and fronting.	52
4.1	The case studies in this chapter, a broad IPA transcription, Wells Lexical Set labels, and IPA transcription of the range of variation.	81
4.2	Coding criteria for /aw/	83
4.3	Token counts of each variant	84
4.4	Coding criteria for /ow/	89
4.5	Token counts of each variant	90
4.6	Dates that /ow/ variants began and stopped fronting, based on δ_{lk}	95
4.7	Maximum /ow/ F2 values, and dates they were reached.	96
4.8	Examples of post coronal /j/ loss in North America	98
4.9	The phonologization of post-coronal /uw/ fronting	98
4.10	Coding criteria for /uw/	99
4.11	Token counts of each /uw/ variant	100
4.12	Comparing the timing of [uw] fronting and the differentiation of [uwL]	105
4.13	Summary of /Vw/ Results	105
5.1	Opaque interaction between /ay/ raising and flapping.	111
5.2	Number of /ay/ observations in each context	115
5.3	Median /ay/ durations by context.	117
5.4	Defining the undergoers and non-undergoers phonetically according to two different precursor hypotheses.	120
5.5	Regression Estimates for word internal /ey/ raising. Reference levels: Decade=1900; FolSeg=C. Model formula: Diag ~ Decade * FolSeg + (FolSeg Speaker) + (Decade Word).	134
5.6	Regression Estimates for word internal /ey/ raising. Reference levels: Decade=1900; FolSeg=/l/. Model formula: Diag ~ Decade * FolSeg + (FolSeg Speaker) + (Decade Word).	135
5.7	Regression Estimates /ey/ raising in inflected vs uninflected words. Reference levels: Decade=1900; Context=/+ed/. Model formula: Diag ~ Decade * Context + (Context Speaker) + (Decade Word).	138

5.8	Regression Estimates /ey/ raising in inflected vs uninflected words. Reference levels: Decade=1900; Context=/+s/. Model formula: Diag ~ Decade * Context + (Context Speaker) + (Decade Word)	138
5.9	Correlation of /uw/, /ow/ and /aw/ across speakers (reported p-values have been Bonferroni corrected within each test statistic).	149
5.10	Comparison of correlation statistics based on speaker means to those based on random effects.	151
5.11	Correlation of /oh/ and /æh/ across speakers.	156
5.12	Comparison of correlation statistics based on speaker means to those based on random effects for /oh/ and /æh/.	157

List of Figures

2.1	The division of North America into /ow/ fronting vs. /ow/ backing regions. From the Atlas of North American English.	17
2.2	The fronting and subsequent backing of /ow/ in Philadelphia	17
2.3	Schematic of the phonology & phonetics grammar.	21
2.4	Mean F1 and F2 values by sex and age in unnormalized Hz	28
2.5	The proposed quantal relationship between changes in articulation and changes in acoustic realizations	31
2.6	The loss of ‘V-to-T’ movement in Early Modern English	33
2.7	Pre-voiceless /ay/ raising	34
2.8	Three distributions differing in standard deviation and kurtosis	35
2.9	An illustration of the systematic relationship between mean, standard deviation, and kurtosis of the mixture of two distributions	36
2.10	Comparing speakers’ means to their standard deviation and kurtosis for /ay0/	37
2.11	Distribution of [ay0] data for the 4 most conservative and 4 most advanced speakers	38
2.12	Simulated speakers at the beginning, middle, and end of the change	39
2.13	The effect of mixing distributions on three different diagnostics	39
2.14	The relationship between normalized F1 mean and kurtosis as observed in speakers, overlaid on the two dimensional density distribution from the mixture simulation	41
2.15	The relationship between normalized F1 mean and standard deviation as observed in speakers, overlaid on the two dimensional density distribution from the mixture simulation	41
3.1	Histogram of how many vowel measurements are drawn from each speaker	49
3.2	Year of interview, and age of speaker	50
3.3	Speakers’ date of birth, and age at time of interview	50
3.4	Relationship between /ay/ raising and year of recording	51
3.5	Relationship between /ay/ raising and speaker’s age	51
3.6	Relationship between /ay/ raising and speaker’s date of birth	51
4.1	Sex differences in the acoustic realization of /ɪ/ in unnormalized F1×F2 space	59
4.2	Distribution of contextual variants of a hypothetical vowel	61
4.3	Independent targets of phonetic implementation produced by phonological differentiation	62
4.4	The effect of coarticulation on shifting productions from intended targets	63
4.5	The interaction of phonological feature spreading and diachronic phonetic change	64

4.6	The interaction of phonetic coarticulation and diachronic phonetic change	64
4.7	The rate of phonetic change in the context of phonological feature spreading	65
4.8	The rate of phonetic change in the context of phonetic coarticulation	67
4.9	The reanalysis of phonetic coarticulation as phonological feature spreading, and its effect on the rate of phonetic change	68
4.10	B-spline basis used in all following models	71
4.11	Weighted b-spline basis, and resulting spline fit	72
4.12	Five randomly generated b-spline curves	72
4.13	The effect of a larger basis on the fit's 'wiggliness'	73
4.14	The full trace for three chains estimating δ_{lk} for $l = 1940$, and the sample approximating the posterior	78
4.15	The Philadelphia Vowel System in the 1970s. From Labov (2001)	80
4.16	/aw/ Trajectory in F1×F2 Space	82
4.17	/aw/ change trajectory	82
4.18	The effect of following nasals on /aw/	83
4.19	The /aw/ variants to be modeled	84
4.20	\hat{R} for the predicted trajectories of /aw/ variants	85
4.21	Predicted trajectories of change for /aw/ variants	86
4.22	Year-to-year differences for variants of /aw/	87
4.23	Rate of change differences from /aw/	87
4.24	/ow/ Trajectory in F1×F2 Space	88
4.25	The effect of following /l/ and a word boundary on /ow/	89
4.26	The /ow/ variants to be modeled	90
4.27	\hat{R} for the predicted trajectories of /ow/ variants	91
4.28	Predicted trajectories of change for /ow/ variants	92
4.29	Year-to-year differences for variants of /ow/	93
4.30	Rate of change differences from [ow]	94
4.31	/uw/ Trajectory in F1×F2 Space	97
4.32	Illustration of the effect of coarticulation on /uw/ from Ohala (1981)	98
4.33	The effect of following /l/ and preceding coronal on /uw/	99
4.34	The /uw/ variants to be modeled	100
4.35	\hat{R} for the predicted trajectories of /uw/ variants	101
4.36	Predicted trajectories of change for /uw/ variants	102
4.37	Year-to-year differences for variants of /uw/	103
4.38	Rate of change differences from [uw]	104
5.1	The effect of following voice and manner on /ay/ height	112
5.2	Effect of following word onset on word-final /ay/	113
5.3	Effect of following <i>to</i> and <i>the</i> on word-final /ay/	113
5.4	Exceptional raising words compared to /ay/ before flapped /t, d/ in all other words	115
5.5	Violin plot representing the distribution of durations of /ay/ before surface and flapped /t/ and /d/	116
5.6	Schematic illustration of the reanalysis of /ay/ raising from being phonetically conditioned to being phonologically conditioned	118
5.7	Nucleus to glide trajectories in Victoria, B	119

5.8	Schematic illustration of the reanalysis of /ay/ raising from being phonetically conditioned to being phonologically conditioned	120
5.9	/ay/ height by date of birth and context	121
5.10	\hat{R} for all parameters in the model	126
5.11	Model estimates of /ay/ F1, faceted by surface vs	127
5.12	Model estimates of /ay/ F1, faceted by /t/ vs /d/	127
5.13	The difference in normalized F1 for /ay/ before flapped /t/ and /d/ from surface /t/ and /d/	128
5.14	The effect of following phonological voice on /ay/ across context	129
5.15	The difference in the effect of voicing between surface and flap contexts	129
5.16	Distribution of /ey/ means by following stop	131
5.17	The effect of following context on word internal /ey/ raising	132
5.18	Trajectory of word internal /ey/	133
5.19	Trajectories of the words <i>day</i> and <i>days</i>	137
5.20	Comparison of word final /ey/ to inflected versions of the same words, as well as /ey/ followed by non-morphological /z/ and /d/	137
5.21	Trajectory of word final /ey/	139
5.22	Trajectory of word final /ey/	140
5.23	The Canadian Chain Shift in reaction to the merger of /o/ and /oh/. Modified from Clarke et al. (1995).	142
5.24	The Pittsburgh Shift in reaction to the merger of /o/ and /oh/. Modified from Labov et al. (2006).	142
5.25	Vowel system of a 66 year old woman from Pittsburgh	143
5.26	The Canadian Chain Shift with /ʌ/ lowering. Modified from Clarke et al. (1995).	144
5.27	Vowel system of a 36 year old woman from Winnipeg	144
5.28	The Canadian Parallel Shift	145
5.29	The Canadian Parallel Shift. Modified from Boberg (2005).	146
5.30	The fronting of back vowels in Philadelphia	147
5.31	The relationship between /aw/ and /ow/ across speakers	148
5.32	The relationship between /uw/ and /ow/ across speakers	148
5.33	The relationship between /uw/ and /aw/ across speakers	149
5.34	Cubic regression spline fits from the generalized additive mixed effects models	152
5.35	The relationship between /aw/ and /ow/ across speakers	152
5.36	The relationship between /uw/ and /ow/ across speakers	153
5.37	The relationship between /uw/ and /aw/ across speakers	153
5.38	The correlation of /oh/ and /æh/ in Philadelphia	155
5.39	/oh/ and /æh/ Random Intercepts	156
5.40	The relationship between /ow/ and /ey/ diphthongization. Figure from Haddican et al. (forthcoming, Figure 2)	160
6.1	Pre-Change Coarticulation, from Ohala (1981)	166
6.2	Phonetic drift based on exemplar simulation	170
6.3	Predicted trajectories of change for /ow/ variants	173
6.4	Year-to-year differences for variants of /ow/	174
6.5	Density distributions of /aw/ and /ow/ along unnormalized F2	175
6.6	Trajectory of word internal /ey/	176

6.7	Quantile Regression over Speaker Means for [ay0]	183
6.8	Interdependent Language Acquisition Tasks	185

Chapter 1

Introduction

In this dissertation, I investigate the interrelationship between phonology and phonetics, specifically with regards to phonetic change. Aided by an unparalleled body of data in the Philadelphia Neighborhood Corpus, I have been able to explore the relative timing of phonological and phonetic influences on phonetic change, and arrived at some novel and interesting results. Specifically, I found that the process of phonologization appears to happen faster, and earlier in the lifespan of phonetic change than previously assumed.

This research is unique in a number of ways. It is the first dissertation to make extensive use of vowel measurements from the Philadelphia Neighborhood Corpus [PNC]. Labov et al. (2013) is the first major publication reporting on results from the PNC, in which we discuss a broad overview of the Northernization of the Philadelphia dialect. We found that those sound changes which Philadelphia shared with the Southeastern super region have been reversing, while those which it shares with Northern dialects have been moving uninterrupted across the 20th century. In this dissertation, I take a more detailed approach to the internal conditioning of many of these changes, with the goal of understanding which conditioning factors can be considered phonetic and which can be considered phonological, and whether a difference between the two can be determined.

Secondly, few other pieces of work investigating phonologization utilize data from language change in progress, while most research utilizing data from language change in progress don't

address themselves to the problem of phonologization. As I make clear throughout the dissertation, language change in progress provides unique insights, and surprising results, that are not readily replicable looking only at the beginning and endpoints of sound change, or only at synchronic experimental results. Those lines of research, exemplified by Ohala (1981), do provide valuable information, but still leave gaps in the model which can only be filled with data from language change in progress. For example, the results from Ohala (1981) argue convincingly that many sound changes result from natural perception errors on the part of listeners. However, it still leaves open the question of *how* perceptual errors lead to sound change. Do the errors accumulate over time within a speaker, or across a speech community? Is the change phonetically abrupt and probabilistic, or phonetically gradual? Do conditioning environments become gradually phonologized, or is phonologization sudden? And at what point in the lifespan of the change does phonologization occur? Data on language change in progress fills in some of these gaps.

In trying to grapple with these issues in sound change, my results are relevant to a broader range of questions about the contested relationship between phonology and phonetics in general. On the one hand, Docherty and Foulkes (2000) and Foulkes et al. (2010) argue that sociophonetic data is best explained using exemplar models of phonetics and phonology (Pierrehumbert, 2002) whereby the primary units of representation are episodic memory traces of the phonetic production of words. In exemplar models, phonological categories emerge out of the statistical regularities of phonetics. On the other hand, the research program of phonetically based phonology (Hayes and Steriade, 2004) pursues the hypothesis that there is not a qualitative difference between phonological and phonetic competence. For example, Flemming (2001, 2004) proposes weighted Optimality Theory constraints which operate over formant transitions, and *n*-ary vowel features.

My results are challenging to both the views that phonological categories are merely codifications of statistical properties of the phonetics, and that there is not a qualitative difference between phonological and phonetic representation and computation. Rather than uncovering an inherent fuzziness to phonological categories, by increasing the volume of data we collect from speakers, the evidence for categorical phonological units has gotten sharper. It appears that cat-

egorical phonological processes which differentiate allophones enter the grammar at the *onset* of conditioned sound changes, rather than as late stage reanalyses. The consequence of this result is that phonological representations cannot simply be the codifications of robust phonetic effects, because at the onset of the change there is no robust effect to be codified. Additionally, the qualitative difference I found between the categorical conditioning of the change, and the fine grained phonetic effects overlaid on the change suggests that there ought to also be a qualitative difference between phonology and phonetics.

The dissertation is laid out as follows. In Chapter 2, I establish my minimal theoretical commitments that I must presuppose in order to make any progress in my data analysis. I first lay out the similarities between typological variation and kinds of sound changes. My point in doing so is to highlight the fact that sound change is necessarily a change in speakers' competence over time, much in the same way that typological variation is differences in speakers' competence across populations. Therefore, the ways that languages can change are strictly constrained by the ways in which speakers' competences can differ. With that in mind, the study of language change is quite clearly the study of linguistic competence. Towards the end of the chapter, I devote a considerable amount of time to describing how phonetic changes occur, in order to assure that I am operating under proper assumptions throughout the rest of the dissertation.

In Chapter 3, I briefly outline the data I use in this dissertation, which is entirely drawn from the Philadelphia Neighborhood Corpus. This chapter is brief so as to avoid considerable overlap with already published descriptions of the PNC (Labov et al., 2013) and of Forced Alignment and Vowel Extraction (Yuan and Liberman, 2008; Evanini, 2009; Labov et al., 2013). I did enhance the data from the output of the FAVE-suite, however, and those enhancements are described there.

Chapter 4 is the first heavy data analysis chapter where I attempt to differentiate between phonological and phonetic conditioning of sound change. The core idea presented in this chapter is that if two variants of a vowel are created in the phonetics, their trajectories over time are yoked together, and are not independent, but if they are created in the phonology, then in principle they can have independent trajectories. The way I evaluate the dependence or independence of vowel variants' diachronic trajectories is to compare their rate of change. This is an extension

of Constant Rate Effect reasoning (Kroch, 1989) frequently utilized in historical syntax. Because the particular changes I examine in Chapter 4 overall have complex diachronic patterns (they moved in one direction, then reversed, as described in Labov et al. (2013)), and because I wanted to investigate the relative timing of phonologization in these changes, I could not rely on standard statistical tools like mixed-effects linear regression. Instead, I construct a custom Bayesian hierarchical model, which is estimated via Hamiltonian Monte Carlo simulation (Stan Development Team, 2012). Of course, a number of complications arise when looking at naturalistic data, but after taking into account possible confounding factors, it appears as if conditioning factors on these vowel shifts fall into two broad categories: those which move in parallel throughout the entire change, and those which were divergent from the outset. At least for these cases, it appears as if categorical phonological conditioning is in place from the outset of the change, and that phonetic conditioning factors were not eventually reanalyzed as being phonological.

In Chapter 5, I examine a number of cases where phonological factors appear to have the greatest explanatory power for both cases where vowel variants have divergent trajectories, and for where multiple vowel categories have parallel trajectories. First, I look at /ay/ and /ey/ raising. These vowels were, for various reasons, imperfect candidates for the rate of change analysis in Chapter 4. However, a close examination of their internal conditioning reveals surprising results. In the case of /ay/, I found that despite the differences in the phonetic contexts of preceding [t] and [d], and flaps corresponding to underlying /t/ and /d/, the raising of pre-voiceless /ay/ took place before underlyingly voiceless contexts, despite their surface realizations. That is, the opaque relationship between raising /ay/ before voiceless consonants and the flapping of /t/ was in place from the very beginning of the change. In the case of /ey/ raising, I find that even though the context of a following /l/ appears to phonetically favor the direction of the change, it does not itself participate. Even other phonetically similar following segments, like /r/ and /w/, condition /ey/ raising, but a following /l/ does not. An explanation for why /l/ would phonetically favor, but not actually condition the change is not forthcoming on strictly phonetic grounds. After looking at /ay/ and /ey/ raising, which exhibit phonologically conditioned divergence, I look at a few cases of parallel shifts. There are two cases of parallel shifts I observe in the PNC. First is

the parallel fronting of /aw/, /ow/ and /uw/, followed by their parallel retraction. Second is the parallel lowering of /æ:/ (tense /æ/) and /ɔ://. I do my best to address the concern voiced by Watt (2000) that these parallel shifts share a social, rather than phonological source, and still find that their parallelism holds.

In Chapter 6, I take the results from the preceding chapters to argue against a model of gradual phonologization. I argue that in each case I examined, evidence of categorical phonologization was observed at the outset of the change, not as a reanalysis later in the change. This result carries with it a number of complications. First, it must be the case that phonetic differences which are small at the beginning of a change correspond to a categorical phonological difference, casting doubt on the hypothesis that phonological categories emerge from reliable statistical properties of the phonetics. Second, it must be the case that new phonological processes are spontaneously hypothesized by language learners. Both of these conclusions may be controversial, so I devote most of Chapter 6 arguing for their plausibility.

In chapter Chapter 7, I provide conclusions, which will largely be a recapitulation of this introduction chapter.

Chapter 2

What is Phonetic Change?

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

2.1 Sound Change and Grammar

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

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

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

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

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

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

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

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

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

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

2.1.1 Phonemic Incidence

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

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

(2.1) North

- a *Don, on*
- ɔ: *Dawn, chocolate*

(2.2) South

- a *Don, chocolate*
- ɔ: *Dawn, on*

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

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

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

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

I also include the linear order of phonological content as falling under this domain of knowledge. A very salient example of cross-dialectal variation in the linear order of phonological material in North America is the difference in *ask* between most White dialects (/æsk/) and African American English (/æks/). This is clearly a difference in lexical knowledge rather than, say, the reflex of different phonotactics, because the difference in /sk/~/ks/ order is restricted to only this lexical item. Similarly, there are examples of lexically sporadic metathesis changes. *The Metathesis Website* (Hume, 2000) provides the example of *chipotle* (an increasingly common word in North America due to the restaurant chain named after the smoke-dried jalapeño), which is sporadically metathesized /tʃipotle > tʃipolte/. There are, of course, many more examples of metathesis in sound change, such as those given by Blevins and Garrett (2004). However, Blevins and Garrett

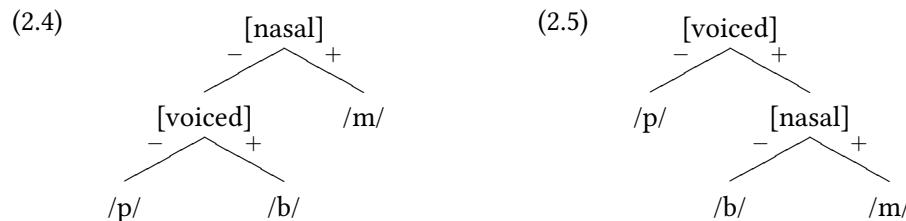
(2004) describe most of their examples of metathesis as fully regular in their outcomes, making it ambiguous as to whether these sound changes progressed as a series of lexically sporadic metatheses, ultimately concluding by spreading across the entire lexicon, or as a the result of a new phonological process or phonotactic being introduced into the grammar elsewhere. This latter option conceptually possible due to productive metathesis processes in synchronic phonological grammars, as Buckley (2011) discusses extensively. Mohanan (1992) and Anttila et al. (2008), for example, describe the following productive alternation for some speakers of Singaporean English.

	Word Final	Intervocalic
(2.3)	<i>lisp</i> [lips] <i>lisping</i> [lispiŋ]	
	<i>crips</i> [krips] <i>crispy</i> [krispi]	
	<i>grasp</i> [gra:ps] <i>grasping</i> [gra:spiŋ]	

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

2.1.2 Systems of Phonological Contrast

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



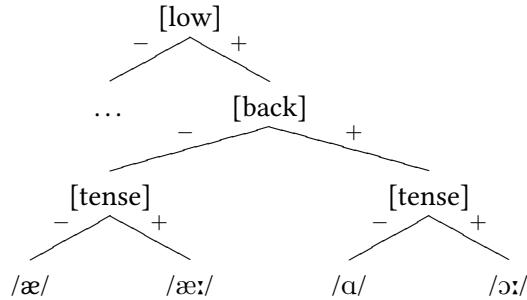
Under the *Contrastivist Hypothesis*, in the language with the contrastive hierarchy in (2.4), /m/ would not participate in, say, voicing assimilation processes, because it is not contrastively specified [+voice]. In a language with the same exact phonemic inventory, but the contrastive hierarchy in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

Recent works supporting the hypothesis that speakers represent contrastive hierarchies such

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

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

(2.6)



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

2.1.3 Presence or Absence of Phonological Processes

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

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

(2.7) “day”: [k]~[g]

(a) *tac* (acc.sg)

(b) *tage* (acc.pl)

(2.8) “strong”: [k]~[k]

(a) *stark* (uninflected)

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

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

(2.9) “day”: [k]~[g]

(a) *tac* (acc.sg)

(b) *tage~tag* (acc.pl)

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

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

(b) stark; starkes > stark; starkes

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

2.1.4 Targets of Phonetic Implementation

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

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

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

The specific targets of phonetic implementation for the same phonological objects can thus vary cross-dialectally and across social groups, meaning that speakers must be able to represent differences in phonetic targets at least as small as the increment of change for women. §2.3.1 will be devoted to arguing that this incrementation is effectively infinitely small since the change

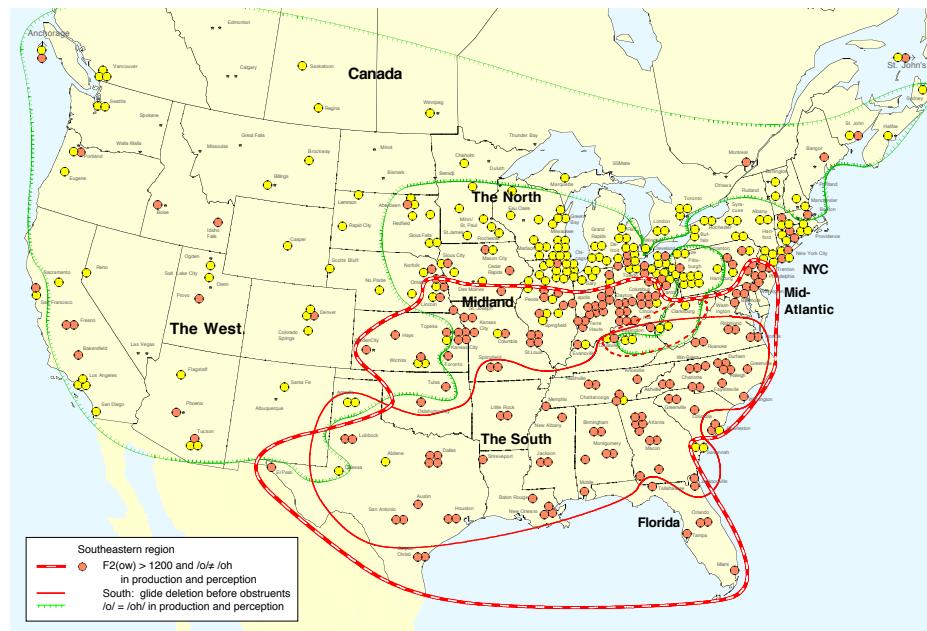


Figure 2.1: The division of North America into /ow/ fronting vs. /ow/ backing regions. From the Atlas of North American English.

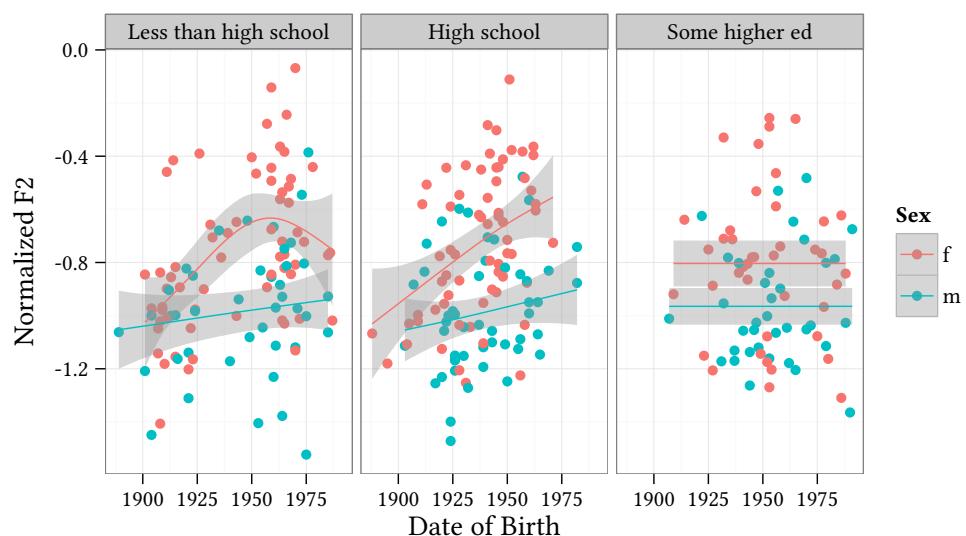


Figure 2.2: The fronting and subsequent backing of /ow/ in Philadelphia.

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

2.1.5 Gestural Phasing and Interpolation

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

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

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

(2.11) *pasta*

pasta > pahta ~ paht^ha ~ pat^ha

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

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

2.1.6 Sound Change and Grammar Summary

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

(2.12) sources of merger

- (a) lexically gradual change in phonemic incidence
- (b) change in the system of contrast
- (c) phonetically gradual change

(2.13) sources metathesis

- (a) lexically gradual change in phonemic incidence
- (b) introduction of a productive phonological process, which outputs metathesis

- (c) gradual change in the alignment of articulatory gestures

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

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

2.2 The Phonology-Phonetics Interface

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

2.2.1 Modular and Feedforward

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

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

2.2.2 The Architecture

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

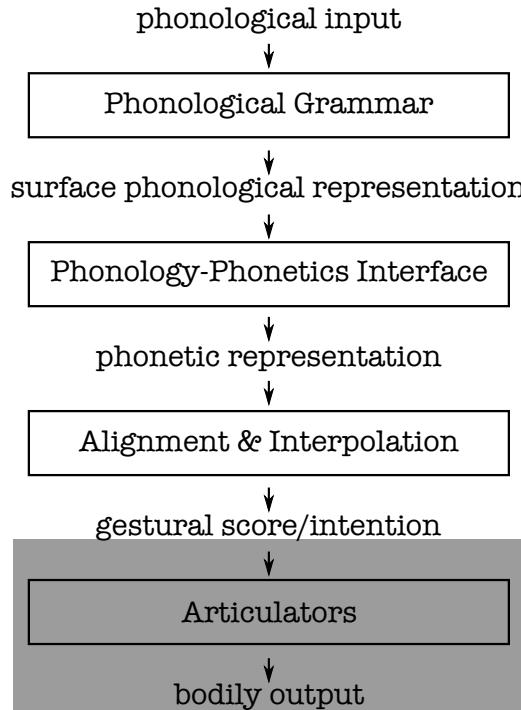


Figure 2.3: Schematic of the phonology & phonetics grammar.

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

Input

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

Phonological Processing

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

Targets of Phonetic Implementation

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

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

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

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

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

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

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

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

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

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

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

(2.16) [+low] \rightsquigarrow 0.1 height

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

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

(2.17) Phonological and phonetic representations are qualitatively different.

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

(2.19) The interface operates over phonological features.

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

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

Phonetic Alignment and Interpolation

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

Universal Phonetics

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

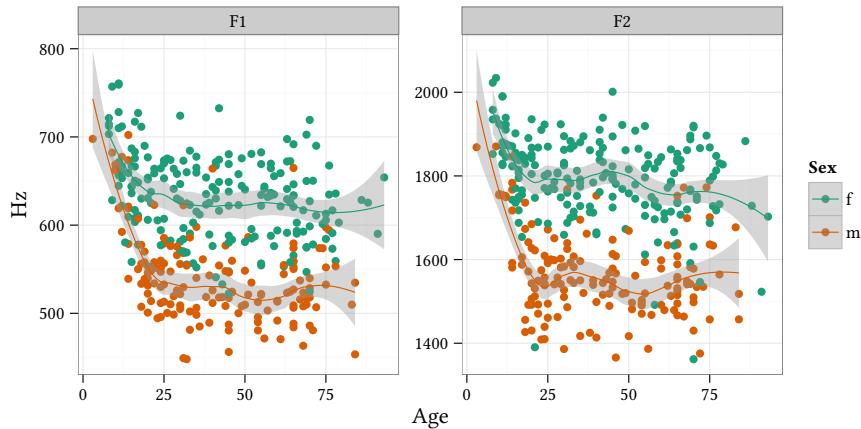


Figure 2.4: Mean F1 and F2 values by sex and age in unnormalized Hz.

2.2.3 Sociolinguistic Variation

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

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

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

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

2.3 Phonetic Change

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

2.3.1 Phonetic Change is Continuous

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

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

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

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

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

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

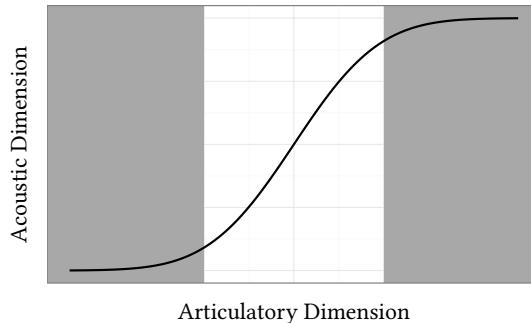


Figure 2.5: The proposed quantal relationship between changes in articulation and changes in acoustic realizations

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

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

Simulating Phonetic Change as Categorical

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

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

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

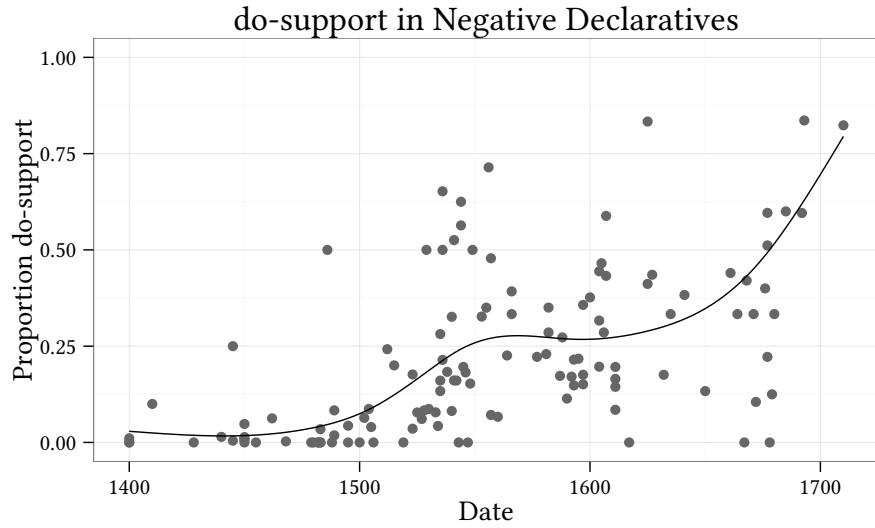


Figure 2.6: The loss of ‘V-to-T’ movement in Early Modern English

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

However, there are other properties of the *distributions* of observations within speakers that can cast some light on whether or not this change progressed as categorical competition between

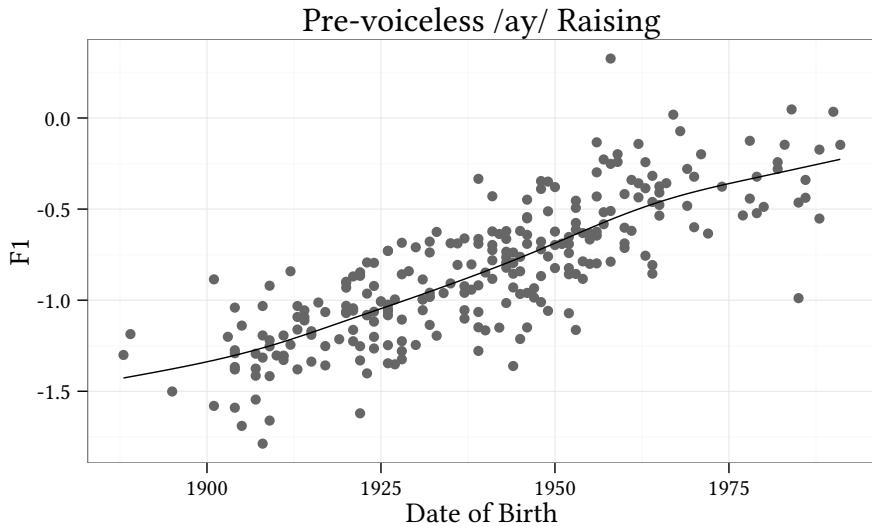


Figure 2.7: Pre-voiceless /ay/ raising.

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

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

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

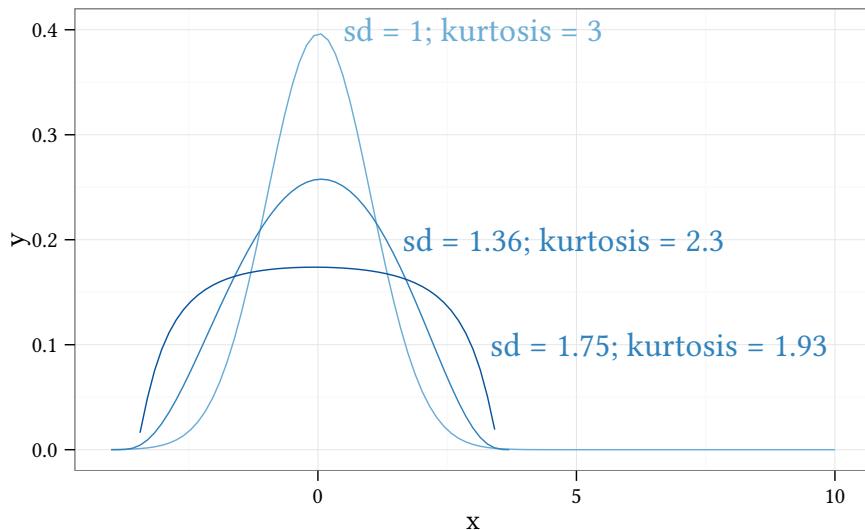


Figure 2.8: Three distributions differing in standard deviation and kurtosis.

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

While the fundamental behavior of these speakers is categorical and probabilistic, given the relative closeness of the phonetic targets for Categories A and B, a researcher would not be able to tell on a token by token basis which category a speaker intended to use in a particular instance.

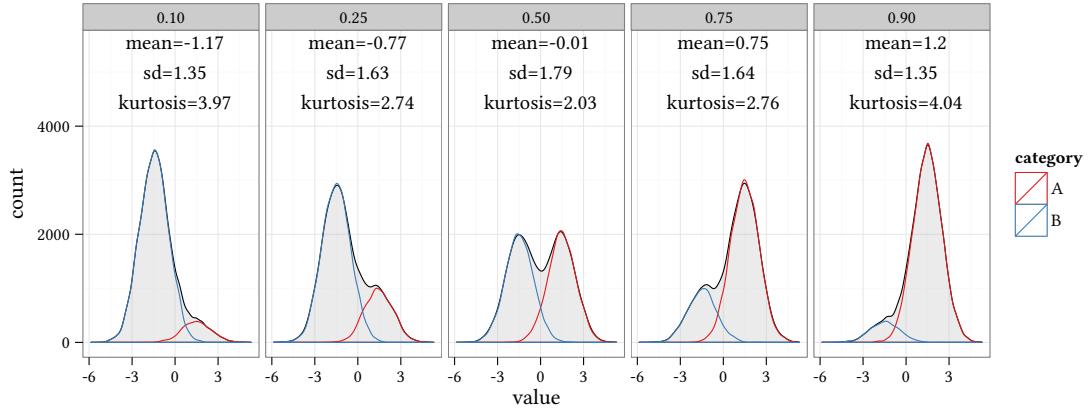


Figure 2.9: An illustration of the systematic relationship between mean, standard deviation, and kurtosis of the mixture of two distributions.

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

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

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

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

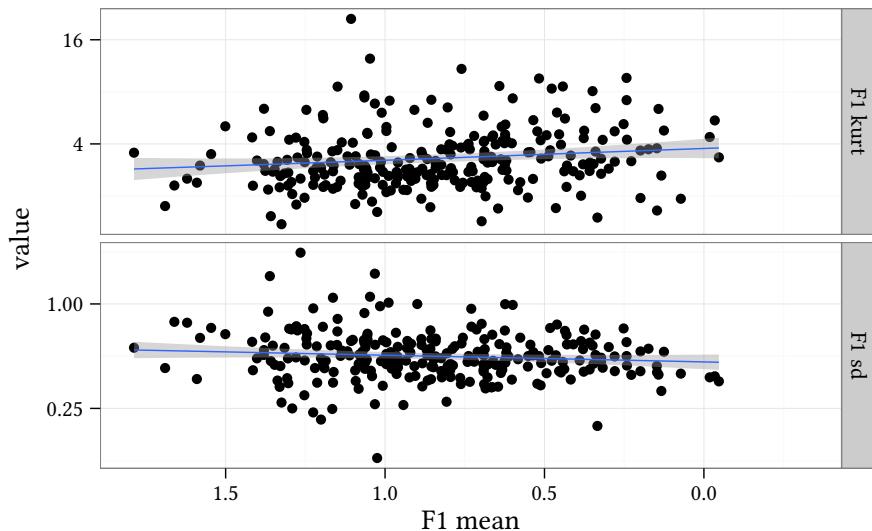


Figure 2.10: Comparing speakers' means to their standard deviation and kurtosis for /ay0/.

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

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

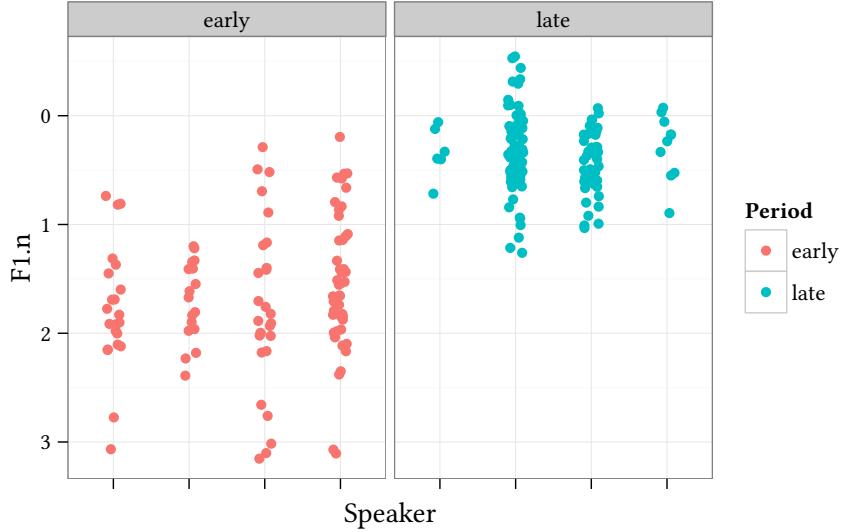


Figure 2.11: Distribution of [ay0] data for the 4 most conservative and 4 most advanced speakers.

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

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

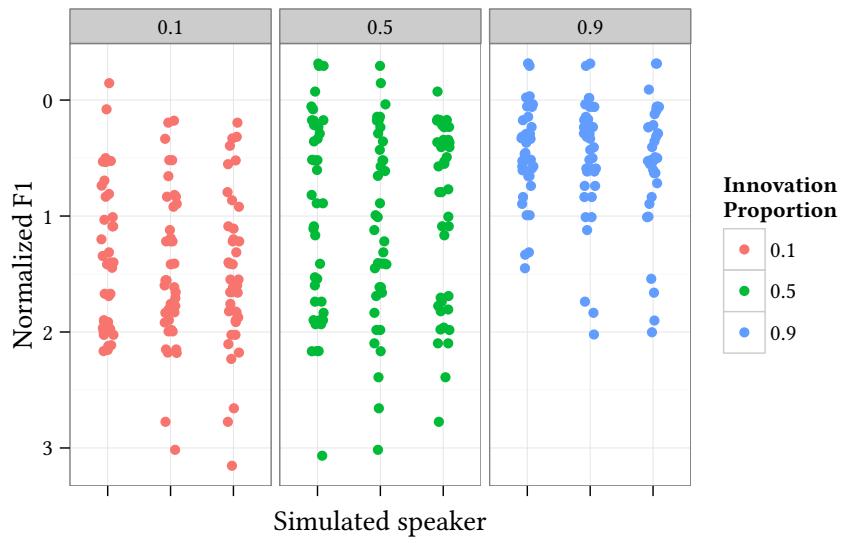


Figure 2.12: Simulated speakers at the beginning, middle, and end of the change.

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

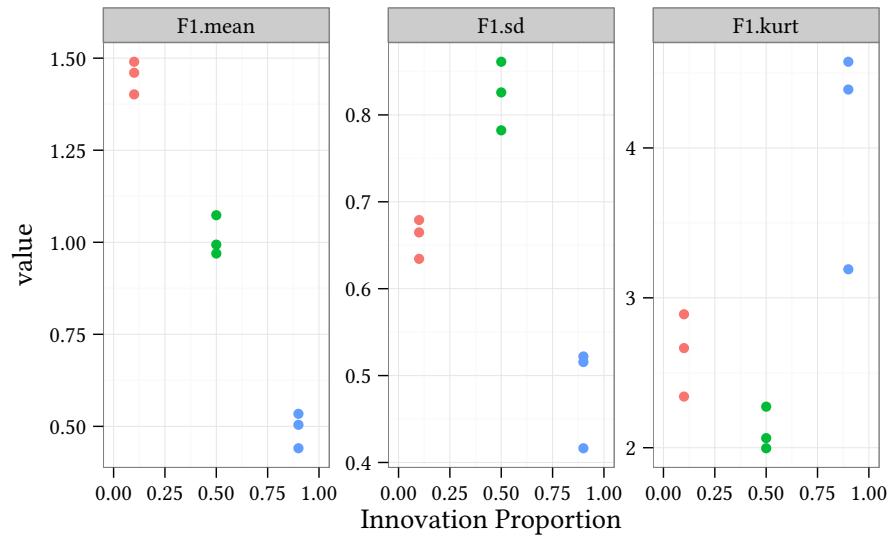


Figure 2.13: The effect of mixing distributions on three different diagnostics: mean, standard deviation, and kurtosis.

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

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

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

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

The conclusion we can draw is that the model of phonetic change whereby /ay/ raised from [aɪ] to [ʌɪ] through categorical variation between these two forms is a poorly fitting one. Rather the fact that both the standard deviation of F1 and its kurtosis remains essentially constant

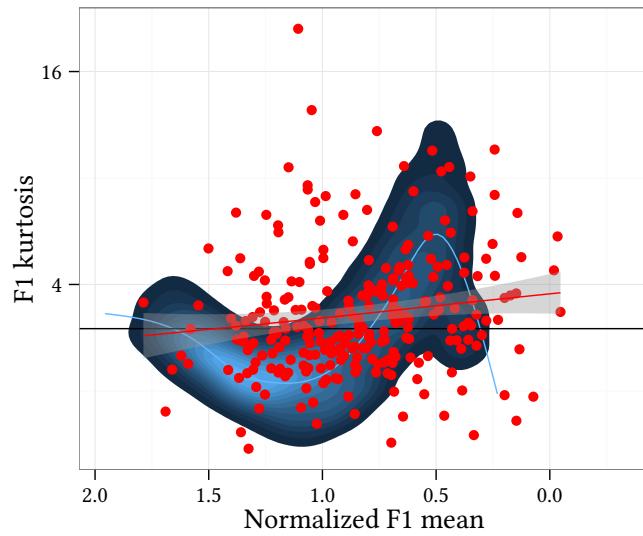


Figure 2.14: The relationship between normalized F1 mean and kurtosis as observed in speakers, overlaid on the two dimensional density distribution from the mixture simulation. Note the y-axis is logarithmic. The horizontal line at kurtosis=3 represents the kurtosis of a normal distribution.

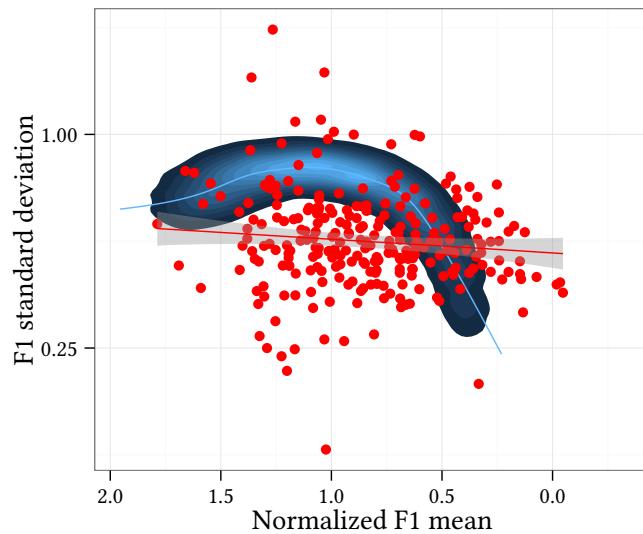


Figure 2.15: The relationship between normalized F1 mean and standard deviation as observed in speakers, overlaid on the two dimensional density distribution from the mixture simulation. Note the y-axis is logarithmic.

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

2.4 Conclusion

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

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

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

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

Chapter 3

The Philadelphia Neighborhood Corpus Data

In this chapter, I'll briefly describe the data used in this dissertation from the Philadelphia Neighborhood Corpus. I'll try not to be overly redundant with descriptions which are already in press (Labov et al., 2013; Evanini, 2009), but I have enriched the data to some extent, which requires some explanation.

3.1 The Philadelphia Neighborhood Corpus

The Philadelphia Neighborhood Corpus [PNC] contains sociolinguistic interviews carried out in Philadelphia between 1972 and 2012 (at the time of this writing). These interviews were carried out as part of coursework for Ling560 ‘The Study of the Speech Community.’ Each year the course was taught (annually from 1972 to 1994, every other year from then on), students formed into research groups, and selected a city block on which to base their study. For more information on Ling560 and the neighborhoods which have been studied, see Labov et al. (2013). The total Ling560 archive contains interviews with 1,107 Philadelphians. Not taking into account the interviews collected in the 2012-2013 academic year, interviews with 379 speakers have been transcribed by undergraduate research assistants, and included in the PNC.

3.1.1 Forced Alignment and Vowel Extraction (FAVE)

The audio recordings and transcriptions of the interviews were then processed by the Forced Alignment and Vowel Extraction (FAVE) suite (Rosenfelder et al., 2011). As the name would suggest, there are two steps to the FAVE analysis. First is forced alignment, which aligns words and phones to the audio. The acoustic models for FAVE come from the Penn Phonetics Lab Forced Aligner [p2fa] (Yuan and Liberman, 2008), with some extra procedures added to account for overlapping speech. With the forced alignment, we can identify where in the audio a particular vowel begins and ends.

The second step is automated vowel formant analysis, an approach first attempted by Evanini (2009). The errors involved in LPC formant analysis are frequently catastrophic, and it was for this reason that the authors of the Atlas of North American English concluded that automated formant analysis was not feasible at the time (Labov et al., 2006). For example, for a vowel like /iy/, in which there is a large distance between F1 and F2, an LPC analysis using 12 poles might erroneously detect a formant between F1 and F2, providing formant estimates with an F2 which is too low. On the other hand, for a vowel like /ɔ/, where F1 and F2 are very close, an LPC analysis using 6 poles might not differentiate F1 and F2, and would return what is actually F3 as F2. In practice, errors like these have been handled by a researcher visually comparing LPC estimates to the spectrogram, and to their personal prior expectations for what the formants of this particular vowel ought to be, adjusting the LPC parameter settings accordingly.

What the FAVE suite does is replace a researcher's prior expectations with quantitative priors from the Atlas of North American English. The vowel class we are trying to measure is given by the forced alignment, meaning that the acoustic data is labeled. Drawing from the ANAE, we can establish certain expectations for the formant measurements we can expect for the specific label. For 4 different LPC parameter settings (6, 8, 10, and 12 poles) we extract the estimates for F1 and F2 frequencies and bandwidths, and compare these to our priors from the Atlas of North American English using the Mahalanobis distance. The LPC parameter setting with the smallest Mahalanobis distance is taken to be the winner. This process is repeated for every vowel in the speaker's interview.

We found that even after we chose LPC parameter settings based on comparison to the Atlas of North American English, there were still a small number of gross errors in the data. We guessed that this may be due to the fact that priors based on the entire ANAE may not be the most appropriate priors for each individual speaker. The most appropriate prior expectation for how a speaker ought to pronounce a vowel is actually how *that speaker* usually pronounces that vowel. Having eliminated most gross errors from the speaker’s vowel measurements through comparison to the ANAE data, we could now generate reasonable speaker specific expectations for each vowel.¹ As a second step, then, FAVE iterates through all of the vowel measurements again, this time comparing the different LPC settings to the speaker specific vowel distributions. This second step, in addition to vowel specific heuristics for selecting a measurement time point,² eliminates almost all gross errors.

For the time being, the FAVE system can only be assured to give high quality results for North American English (because of the reliance on ANAE priors), and only for an aligned corpus using the CMU dictionary transcriptions (which is what p2fa and FAVE-align use). However, extending the method to any given dialect or language is conceptually trivial. First, a certain number of high quality hand measurements for each vowel in the dialect needs to be collected. The ANAE is a large database, but the sample size necessary to establish the first pass priors need not be as large. Even relatively small numbers of measurements, say 10 to 20 per vowel, ought to be sufficient, since the goal of the priors is not to provide an overly precise estimate for each vowel, but rather just to weed out the grossest errors. With these priors collected, the implementation of FAVE would need to be changed to not make specific reference to the CMU labels. Of course, FAVE-extract is dependent on having an alignment, which for now is based on the models from p2fa. However, for other dialects or languages, there are other trainable aligners available, like the Prosodylab-Aligner (Gorman et al., 2011).

¹This was my own substantive contribution to the FAVE suite.

²These were explored and implemented by Ingrid Rosenfelder.

3.2 Enrichment of Contextual Information

FAVE-extract provides two different kinds of output file types: the Plotnik file formant (Labov, 2006a), and a tab delimited file. For the purpose of my dissertation, the data available in these outputs was not entirely sufficient. With regards to the contextual coding, they indicate the place, manner and voicing of the following segment, the nature of the syllable coda, and how many following syllables in the word, and the quality of the preceding segment. The actual label for the preceding or following segment is not included, nor the transcription of the word, or any contextual information across word boundaries. Some of this information is crucial for my analyses, so I enriched the information available from the original FAVE-extract output. Using the time stamp of the measurement point, I scanned the Praat TextGrids which served as the input for FAVE-extract and tried to identify the vowel that corresponded to a particular measurement. This step was complicated by the fact that some vowels were not measured within the boundaries provided by the FAVE-align, but rather within the boundaries of the preceding segment. Any vowel which could not be programmatically located in a TextGrid were discarded. In addition, one speaker's TextGrid could not be located in the corpus, and their data has also been excluded from this dissertation.

Once locating the correct vowel in the TextGrid, I extracted the following information for the vowel.

- (3.1) the full CMU transcription for the word the vowel is located in.
- (3.2) the preceding segment, disregarding word boundaries.
- (3.3) the following two segments, disregarding word boundaries.
- (3.4) the location of the vowel in the word, coded as
 - (a) word initial
 - (b) word final
 - (c) coextensive with word boundaries (e.g. *I*)
 - (d) word internal

These additional pieces of information were crucial for most analyses in this dissertation. For example, for word final vowels, knowing the following segment was crucial for investigating whether the conditions on certain phonetic changes applied at the phrase or word level. Also, with the full CMU transcription of the word, I was able to apply simple syllabification algorithms which allowed me to, for example, compare open and closed syllables on the conditioning of /ey/, and to identify which following /t/ and /d/ were flapped when looking at /ay/.

3.3 Total Data Count

The Ling560 fieldworkers have visited a relatively racially diverse set of neighborhoods. However, the vast majority of speakers included in the PNC so far are of White European descent. It would be a mistake to treat data drawn from African American Philadelphians and White Philadelphians as being drawn from one unified speech community. The two social groups clearly form separate, but mutually influencing, speech communities (Labov et al., 1986). In fact, Henderson (2001) found that listeners could correctly identify White and African American Philadelphians' race simply from a recording of them counting from 1 to 20. Despite the facts that the mutual influence of these two dialects on each other is so interesting, and that the White Philadelphian dialect is spoken now by a numerical minority of all Philadelphians, the nature of the data available at the moment constrains me to look exclusively at White speakers.

Taking into account that I will only be examining the data from White Philadelphians, that one speaker had to be excluded because I could not locate their TextGrid, and that some vowel measurements had to be excluded because the vowel could not be programmatically located in their TextGrid, I will be working with 735,408 vowel measurements from 308 speakers. Figure 3.1 plots a histogram of how many vowel measurements are available from each speaker.

3.4 Normalization

All of the data were normalized to formant intrinsic z-scores (i.e. Lobanov Normalization) (Adank et al., 2004). In this dissertation, I will be using the z-score measure directly, rather than rescaling

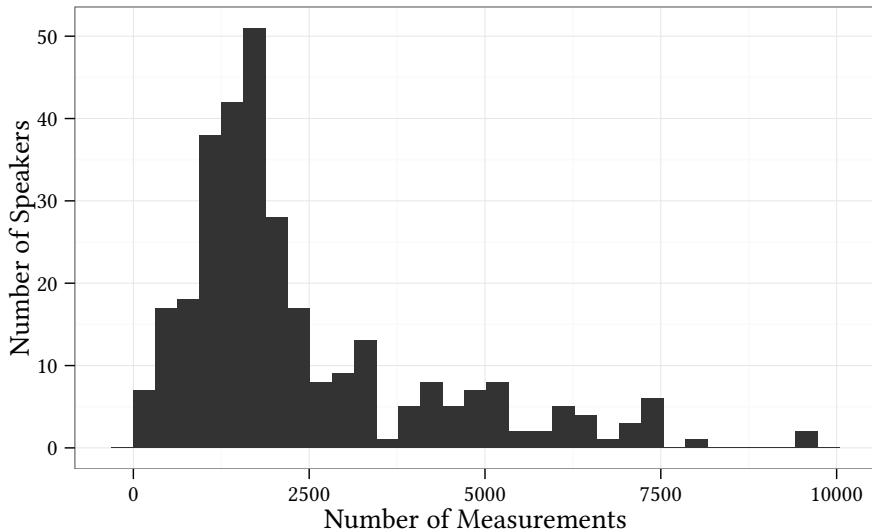


Figure 3.1: Histogram of how many vowel measurements are drawn from each speaker.

it to a hertz-like measure.

3.5 Choice of Time Dimension

There are a number of different possible time dimensions available from the corpus (see Sankoff (2006) for an overview of real and apparent time). I could, for example, use a strictly real time measure, and evaluate the phonetic changes I investigate against their year of interview. There are also two different apparent time measures available, speakers' Age and Date of Birth. Figure 3.2 plots the year of interview and the age of the speaker, while Figure 3.3 plots the date of birth of the speaker, and their age at the time of the interview. What should be clear from Figure 3.3 is that any result obtained using a speaker's date of birth is going to be very similar if the speaker's age was used instead. The high correlation between speaker's age and date of birth is simply due to the facts of human lifespan, and that the fieldwork has covered 40 years.

However, it is possible to compare statistical models that use each kind of time dimension to see which has the best predictive power. Labov et al. (2013) did this by comparing the r^2 of

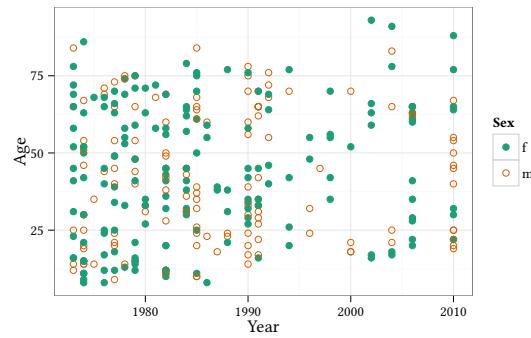


Figure 3.2: Year of interview, and age of speaker.

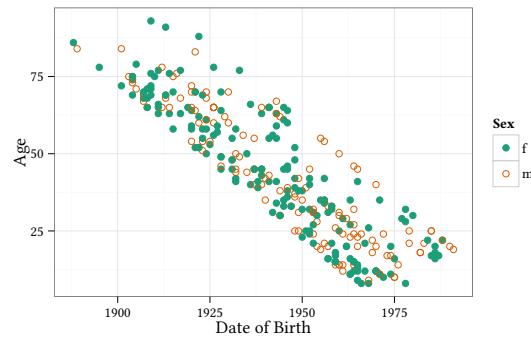


Figure 3.3: Speakers' date of birth, and age at time of interview

each, and I will briefly replicate that analysis here. Figures 3.4 to 3.6 plot the relationship between the normalized F1 of pre-voiceless /ay/ and the three possible diachronic dimensions (year of interview, at at interview, and date of birth).

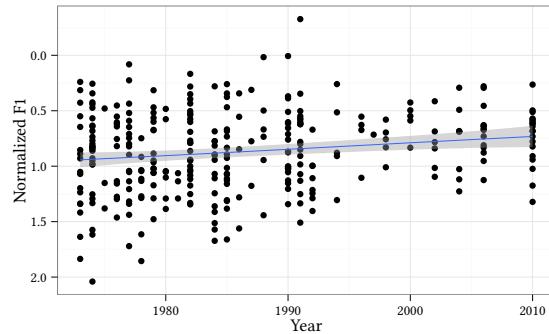


Figure 3.4: Relationship between /ay/ raising and year of recording.

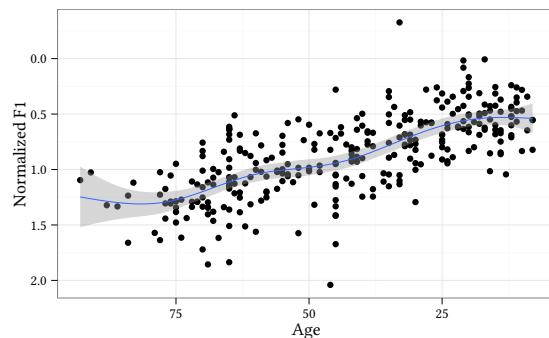


Figure 3.5: Relationship between /ay/ raising and speaker's age.

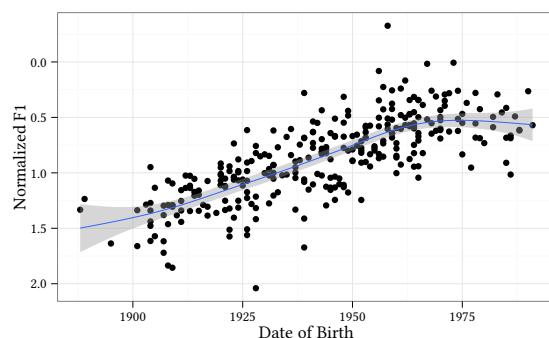


Figure 3.6: Relationship between /ay/ raising and speaker's date of birth.

I fit three generalized additive models to predict the F1 of pre-voiceless /ay/ using cubic regression splines for each sex.³ Table 3.1 displays the r^2 and the Akaike Information Criterion [AIC] for each model. The model predicting pre-voiceless /ay/ F1 using speakers' date of birth has the highest r^2 and the lowest AIC, suggesting that this ought to be the preferred model.

Predictor	r^2	AIC
Year	0.03	247
Age	0.49	53
Date of Birth	0.59	-11

Table 3.1: Model comparisons for using different time dimensions to predict pre-voiceless /ay/ height.

Pre-voiceless /ay/ raising exhibits one of the two patterns of change in Philadelphia that Labov et al. (2013) identified (linear incrementation). Just to make sure that date of birth is also the best diachronic dimension for the other pattern of change (reversal), I fit three models predicting the fronting and raising of /aw/ using year of the recording, speakers' age, and speakers' date of birth. The r^2 and AIC for these models are displayed in Table 3.2. The model using date of birth again has the highest r^2 and lowest AIC, suggesting that for the changes which reversed course, date of birth is also the best diachronic dimension to use. The fact that the r^2 for the best /aw/ is much smaller than the r^2 of the best /ay/ model is probably due to the fact that /aw/ is more highly differentiated along social dimensions, as Labov et al. (2013) found when they took into account speakers' level of education.

Predictor	r^2	AIC
Year	0	454
Age	0.11	423
Date of Birth	0.13	417

Table 3.2: Model comparisons for using different time dimensions to predict /aw/ raising and fronting.

Given that date of birth has the best predictive power for both /ay/ and /aw/, which themselves exemplify the two major patterns of change I investigate in this dissertation, I'll be using date of

³The formula was `gam(F1.n ~ s(X, bs = "cs", by = Sex))`.

birth as the diachronic dimension throughout the dissertation.

Chapter 4

The Rate of Phonetic Change

When examining the effect that one speech segment has on an adjacent segment, there is a persistent problem involved in trying to determine whether that effect should be attributed to phonetic coarticulation, or to a phonological process, especially since the two can appear to be so similar (a so-called “duplication problem”, (Ohala, 1990; Cohn, 2007)). This problem is compounded when these effects are spread out across generational time. The difficulty in distinguishing between phonetic coarticulation and phonological processes with synchronic data, among other things, has led some to propose a much more phonetics-like model of phonology, where the phonology operates over much smaller granular primitives (e.g. Flemming, 2001), and where gradient phonetic realizations are subject to phonological considerations, like contrast.

Meanwhile, an increasing volume of research makes appeals to language change to explain phonological processes, and the apparent naturalness of phonology. Evolutionary Phonology, as proposed by Blevins (2004), is a good example. Blevins states the central premise of Evolutionary Phonology this way:

Principled diachronic explanations for sound patterns have priority over competing synchronic explanations unless independent evidence demonstrates, beyond reasonable doubt, that a synchronic account is warranted. (Blevins, 2004, p 23.)

A key problem for lines of research like this one is that few utilize evidence from language change *in progress* to support their arguments. Most of the argumentation in Blevins (2004), for

example, is based on comparative reconstructions, which provide us with proto forms A, and daughter forms B, C, and D, thus indicating 3 different sound changes: A → B, A → D, and A → C. Blevins (2004) follows up the identification of sound changes like these with argumentation for the phonetic naturalness of each change, and how the change may have occurred given this phonetic naturalness. Blevins proposes some possible mechanisms for sound change (CHANGE, CHANCE, CHOICE), but these mechanisms are supported only by the conceptual plausibility that they may have generated changes A → B, etc., not by direct evidence of these mechanisms at work in a sound change in progress.

Another good example of an appeal to sound change that lacks support from change in progress is Ohala (1990). In that work, Ohala examines the phenomenon of consonantal place assimilation. First, he identifies $C_1C_2 \rightarrow C_2C_2$ as a common sound change.

	Latin	Italian
(4.1)	<i>scriptu</i>	<i>> scritto</i>
	<i>nocte</i>	<i>> notte</i>

Then, he reports results of experiments where various manipulations to non-word sequences like [apta] can affect whether English listeners report hearing [apta], [atta], or [appa]. Unsurprisingly, subjects in the studies were much more likely to misperceive [apta] as [atta] (93%) than as [appa] (7%). The inference that Ohala (1990) makes is that these experimental subjects were, in some sense, recreating a sound change of the type in (4.1). However, even when taking together the experimental results with matching attested sound changes, the way in which the change took place remains underdetermined. The change from $C_1C_2 \rightarrow C_2C_2$ could have been lexically gradual, slowly diffusing through the lexicon, or it could have been lexically abrupt. It could have started in one context (say kt > tt), then spread to other contexts, or it could have affected all contexts simultaneously. In the case of consonantal place of articulation, it's unlikely that this change would have been phonetically gradual, but in the case of post-coronal [u] fronting, another example from Ohala (1981), it's an open question whether it would progress in a phonetically gradual way, or abruptly. The fact that the way language changes like $C_1C_2 \rightarrow C_2C_2$ are underdetermined by experimental work like Ohala (1990) is not just a descriptive gap, but an explanatory one. As

I argued in Chapter 2, the way in which language change progresses is determined by what part of speakers' linguistic competence is changing, meaning one's theory of linguistic competence defines possible paths of language changes, and vice versa. A solid result coming out of Ohala (1981, 1990) is that there appears to be a relationship between the kind of persistent errors listeners make and the outcomes of sound change, but I would argue that is a new fact to be explained, not an explanation itself.

Simulation of language change has also become an increasingly common tool for researchers interested in language change. Unfortunately, the success of these simulations is usually judged by comparing the initial and final states of the simulation to the initial and final states of attested sound changes, rather than by comparing the dynamics of change in the simulation to the dynamics of a known change in progress. For example, Boersma and Hamann (2008) try to model the fact that speech sounds tend to be maximally dispersed along acoustic dimensions by using agent based simulations of cross-generational language acquisition with bidirectional constraint grammars. Their results are interesting and compelling, but their conclusion that their model is a success is based on the fact that it produced maximally dispersed distributions, *not* that they produced realistic patterns of language change when compared to other language changes in progress.

I am proposing that theoretical models like the ones I've just mentioned must compare their predictions about the dynamics of language change to the dynamics of actual language changes in progress in order to claim definitive empirical support. Fortunately, there is a well established field of inquiry into the dynamics of language change in progress, Quantitative Sociolinguistics, with similarly well established methodologies for the study of language change in progress (e.g. Labov, 1994, ch. 3, 4). In order to compare the results of simulation and experiment to language changes in progress, it is crucial to hash out exactly what patterns of language change we ought to see based on different theories of phonology and phonetics, which is one of the partial goals of this dissertation.

The goal of this chapter is to both introduce a novel technique for distinguishing between phonetic and phonological influences on phonetic change, and to establish some basic facts about

the dynamics of sound changes which are subject to some kind of conditioning factors: comparing the rate of inter-generational sound change of vowels across different linguistic contexts. While these basic facts will be of considerable intrinsic interest to theories of language change, I believe I've made it clear that they will also be of considerable interest to phonological theory more broadly construed.

4.1 Phonetic Coarticulation vs Phonological Differentiation

Strycharczuk (2012, ch. 2) outlines a number of ways that researchers have attempted to distinguish between phonological processes and phonetic coarticulation.

- (4.2) Compare segments which are ambiguous between phonetic coarticulation and phonological assimilation to segments which are unambiguous. e.g. compare intervocalic /s/, which may undergo either categorical voicing assimilation or phonetic voicing coarticulation, to phonemic /z/.
- (4.3) Examine the coarticulatory effect over the duration of the segment. A phonetic cline, with its highest point adjacent to the coarticulatory source is indicative of phonetic coarticulation, while a phonetic plateau across the entire duration of the segment is indicative of a phonological process (Cohn, 1993).
- (4.4) Estimate the bimodality of the phonetic distribution of the ambiguous segments, with the hypothesis that strong bimodality is indicative of a phonological distinction.
- (4.5) Examine the coarticulatory effect's sensitivity to speech rate. The hypothesis is that phonetic coarticulation should be sensitive speech rate, but phonological assimilation should not be.

Both (4.2) and (4.3) appear to me to be reasonable approaches to the problem, but unfortunately not universally applicable. None of the cases studies I will be investigating involve neutralization, which is key for (4.2), comparing the phonetics of derived segments to underlying segments. For example, I will be looking at the effect of nasals on the /aw/ diphthong. The most

conservative realization for the nucleus of this diphthong is [æ], when followed by oral segments. However, even the most conservative realizations of the /aw/ nucleus are considerably fronter and higher when followed by nasal segments, [æ_↑~e̥]. I'm unable to utilize (4.2), because pre-nasal /aw/ isn't neutralized to a different segment which appears independently, so I have no unambiguously phonological form of [æv] to compare pre-nasal /aw/ to.

The next option of comparing phonetic clines to plateaus (4.3) is also difficult to bring to bear on the case studies at hand. To begin with, the Philadelphia Neighborhood Corpus, in the form I've had available for this dissertation, only contained point measurements for the nuclei of diphthongs. However, it is even difficult conceptually to determine what would constitute a cline, and what would constitute a plateau in the cases I will be looking at. Using the example of /aw/ again, its raising and fronting when adjacent to a nasal is undoubtedly related to nasality in some way. However, the dimension along which the effect of the following nasal plays out is in vowel height and frontness, which are only indirectly related to nasality. Moreover, /aw/ is an intrinsically dynamic speech segment with two targets. Determining whether the effect which fronts and raises the nucleus of /aw/ is somehow stronger in the glide, or whether it's a constant effect throughout the entire diphthong would be a complicated exercise indeed.

The remaining two options, examining bimodality of the distributions (4.4) and determining speech rate effects (4.5) could be feasibly applied to the cases I'm examining, but there are good reasons to call the diagnostic validity of these approaches into question. To begin with bimodality, it is trivial to come up with examples of bimodal distributions which clearly don't correspond to phonological differences. Figure 4.1 plots the distribution of mean F1 and F2 measurements for /ɪ/ for all speakers in the PNC. The distribution of /ɪ/ is strongly bimodal, but this bimodality is due to the sex of the speaker, since Figure 4.1 is displaying unnormalized data. There is no reason to believe that men and women have fundamentally different phonological representations or even different intended phonetic implementations for /ɪ/. Rather, men and women clearly have the same targets of phonetic implementation for the same phonological object, and those targets have then been filtered through phonetic contingencies (the sex linked differences in vocal tract length).

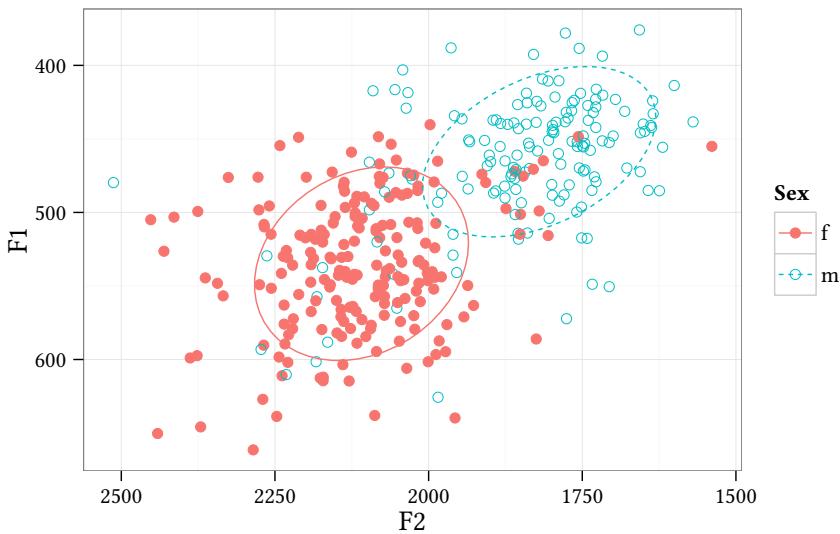


Figure 4.1: Sex differences in the acoustic realization of /ɪ/ in unnormalized F1×F2 space.

The inter-speaker effect of vocal tract length on the realization of vowels is an extreme case of what I will henceforth be referring to as a “phonetic effect.” However, at the moment there is no theory of what the upper limit of intra-speaker phonetic effects due to coarticulation ought to be, especially if the degree of articulatory overlap is a language specific property, per the discussion in §2.1.5. Another case study in this chapter will be on the effect of a following /l/ on preceding /ow/ and /uw/. As /l/ in Philadelphia is frequently much more glide-like, especially in coda positions (Ash, 1982), with its primary place of articulation being dorsal, it is conceivable that it may have a considerable coarticulatory effect on /ow/ such that a bimodal distribution of [ow]~[owl] is the product. This is doubly so if the phonetic alignment constraints for the Philadelphia dialect allow for substantial gestural overlap of the /ow/ vowel and the dorsal /l/ gesture. Given these facts, it is not strictly necessary that strongly bimodal distributions are indicative of phonologically distinct targets.

Furthermore, there is also no theory for what the *lower* limit of phonetic difference is for two phonologically distinct targets. For example, Labov and Baranowski (2006) note that in the Inland North dialect region, the lowering and backing of /ɛ/ and the fronting of /a/ has led to considerable

overlap between these vowels for many speakers, without resulting in merger. They argue that an average duration difference of 50ms is sufficient to maintain and signal the phonemic difference in this case. This is a relatively small difference. For comparison, I calculated the Median Absolute Deviation⁴ for the duration of /a/ for all speakers in the Philadelphia Neighborhood corpus. The median MAD across speakers is 45ms. While it is difficult to make direct comparisons between these two studies due to the drastic differences in the dialects, the fact remains that the size of between category differences in the Inland North is about the same size as the within category variation in Philadelphia.

So, it is both the case that strong phonetic bimodality is not necessarily an indicator of phonological differentiation, and the absence of strong phonetic bimodality is not necessarily an indicator of the absence of phonological differentiation. As such, I will not be utilizing bimodality as a diagnostic for distinguishing between phonetic and phonological effects.

The fourth option, determining whether the effect of one segment on another is sensitive to speech rate (4.5) would be possible to implement with the PNC data. However, the operating assumption behind this method that phonological processes should not be sensitive to speech rate does not stand up to the results of sociolinguistic research. The concept of a variable phonological rule was first introduced by Weinreich et al. (1968), and since then, variable linguistic processes of all sorts have been found to be sensitive to both grammatical and extra-grammatical variables, like speaking style.

Using the case of /ow/ followed by /l/ to make this argumentation concrete, we could imagine that there is a variable phonological process which spreads some additional dorsal feature from /l/ to /ow/, producing a phonetically fully back [o:]. This phonological process could be close to categorical at extremely fast speech rates, but as speech rate slows, its probability of application falls off. When /l/ doesn't spread its phonological features to /ow/, however, it might still be phonetically coarticulated with /ow/, an effect which itself might decrease as speech rate slows even further. The resulting data would appear to show a gradually decreasing effect of /l/ on /ow/ as speech rate decreases, and we would miss the generalization of a phonological process at work

⁴The MAD is calculated by first calculating the distance of all data points from the sample median, then taking the median of their absolute values. i.e. $\text{median}(|x_i - \text{median}(x)|)$

if we were to interpret this to mean the effect of /l/ on /ow/ is purely coarticulatory.

4.1.1 Phonological vs Phonetic Processes in Sound Change

I would like to bring evidence from sound change to bear on the question of whether the influence of one segment on another is due to phonetic coarticulation of phonological differentiation. Let's assume that we are analyzing some hypothetical vowel, /V/, which appears in two different segmental contexts, /__x/ and /__y/. The distributions of [V_x] and [V_y] in F1×F2 space are given in Figure 4.2.

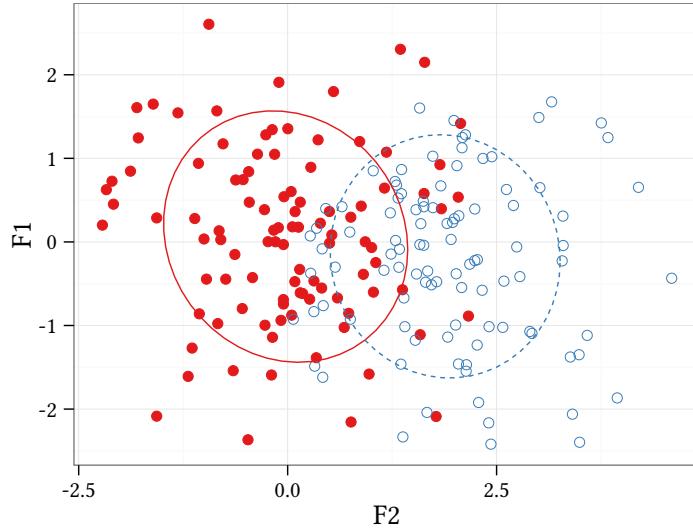


Figure 4.2: Distribution of contextual variants of a hypothetical vowel

There are two distinct ways in which the data in Figure 4.2 could have been generated. First, /y/ could have spread some feature f onto V, creating a featurally distinct, thus phonologically distinct allophone of /V/.

$$(4.6) \quad \text{V} \rightarrow \text{V}_f / __ y$$

As phonologically distinct objects, [V] and [V_f] can have independent targets for phonetic implementation. The target of implementation for [V_f], in this case, happens to be further back along F2. In Figure 4.3, the two independent targets for [V] and [V_f] are represented as two larger points in the centers of their respective distributions.

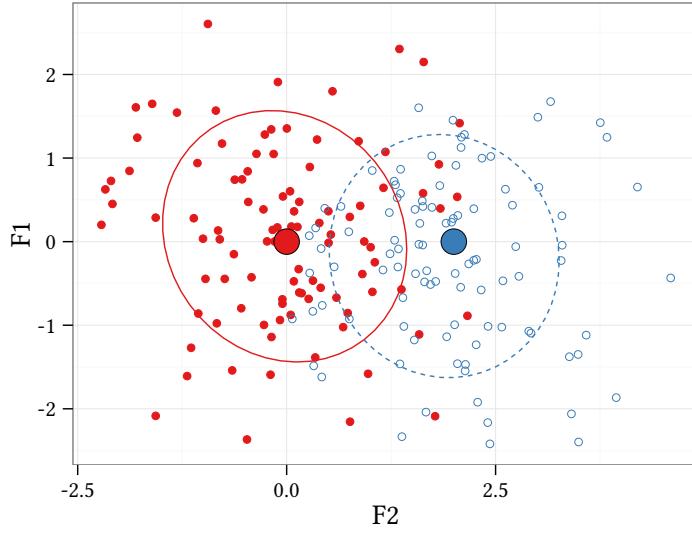


Figure 4.3: Independent targets of phonetic implementation produced by phonological differentiation.

Alternatively, there could be no phonological process involved here at all. Instead, the mapping from phonological representations to phonetic targets could produce only one target, that for [V_x]. However, segment [y] exerts a large coarticulatory pressure on [V], pulling the actual productions of [V_y] back from their intended target. This coarticulatory shift is represented by the arrow in Figure 4.4. The distribution for [V_y] does not have a larger point at the center of its distribution in Figure 4.4 in order to indicate that it does not have its own independent target for phonetic implementation.

As I have argued above, it is not possible to distinguish between these two scenarios given the most common methodologies, nor by just eyeballing the data. However, we should expect to see different patterns in *diachronic* change depending on which process is operating. The key difference is that in the case of phonological feature spreading, [V] and [V_f] have independent targets of phonetic implementation, while in the case of phonetic coarticulation, the realization of [V_y] is yoked to [V_x]. Thus, it should be possible for these contextual variants of /V/ to have separate diachronic trajectories *only* in the case of phonological feature spreading, while in the

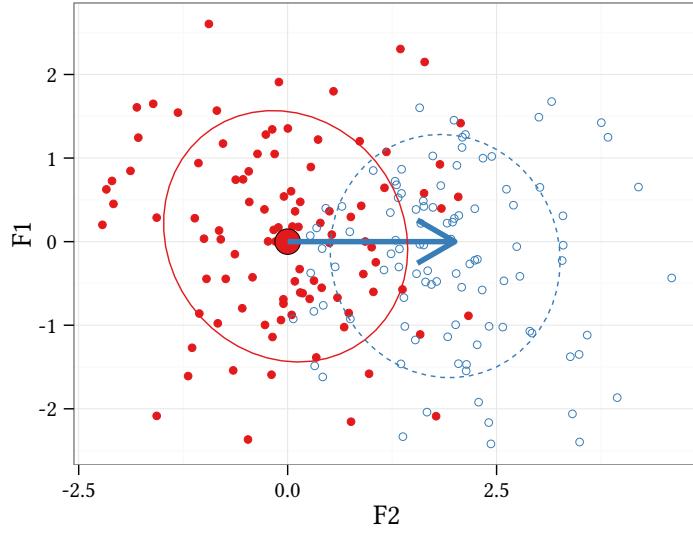


Figure 4.4: The effect of coarticulation on shifting productions from intended targets.

phonetic coarticulation case, the realization of one variant should be yoked to the diachronic trajectory of the other.

Figure 4.5 illustrates the interaction between phonological feature spreading and diachronic phonetic change. The data in this figure represents a shift in one generation from Figure 4.3 where the target of [V] has shifted frontwards along F2, but the target of $[V_f]$ has remained stable. The target for [V] from the previous generation is represented as a large faint point. The important point is that [V] has shifted independently from $[V_f]$, which contrasts sharply with Figure 4.6.

Figure 4.6 represents the interaction of phonetic coarticulation and diachronic phonetic change. Again, the target for $[V_x]$ has shifted frontwards along F2, but because the realization of $[V_y]$ is the product of a coarticulatory shift, which has remained constant, $[V_y]$ has also shifted frontwards along F2.

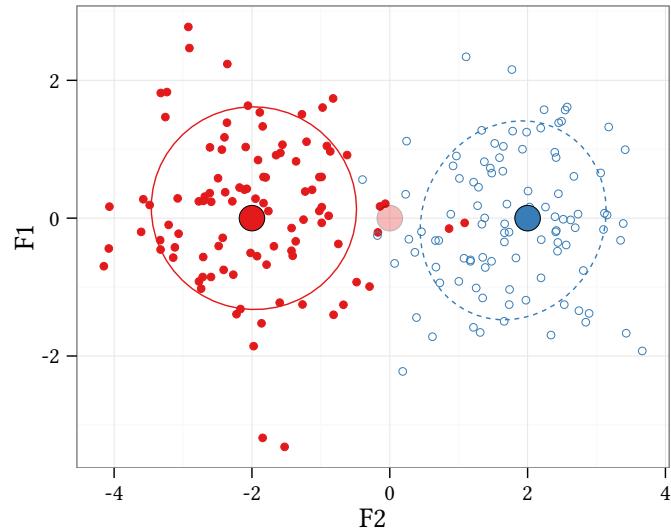


Figure 4.5: The interaction of phonological feature spreading and diachronic phonetic change.

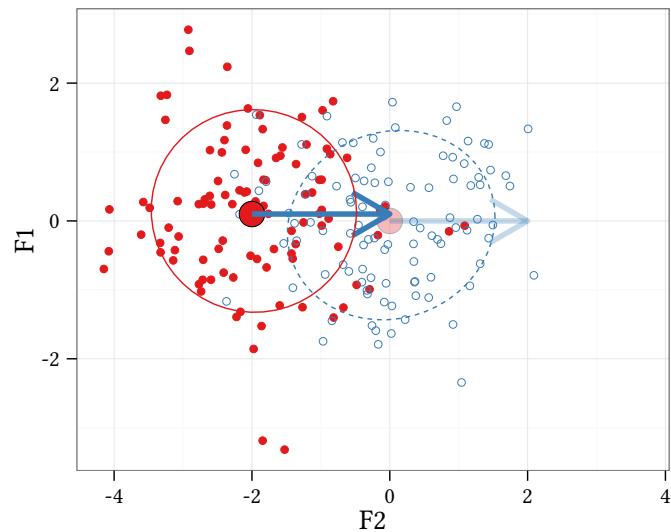


Figure 4.6: The interaction of phonetic coarticulation and diachronic phonetic change.

4.2 The Rate of Language Change

In this section, I'll be fleshing out more completely the way in which phonological feature spreading and phonetic coarticulation produce different predicted dynamics of sound change. Figure 4.5 illustrates the expected difference between two generations when phonetic change interacts with phonological feature spreading. $[V]$ moves frontwards along F2, leaving $[V_f]$ behind. Figure 4.7 presents a finer grained illustration of this effect over age cohorts. The top facet of Figure 4.7 illustrates the target for $[V]$ moving along F2 from 0 to 2 along a classic S-shaped trajectory. The phonetic target for $[V_f]$, on the other hand, remains constant at -2. The bottom facet of Figure 4.7 represents the year-to-year change in F2 for $[V]$ and $[V_f]$.

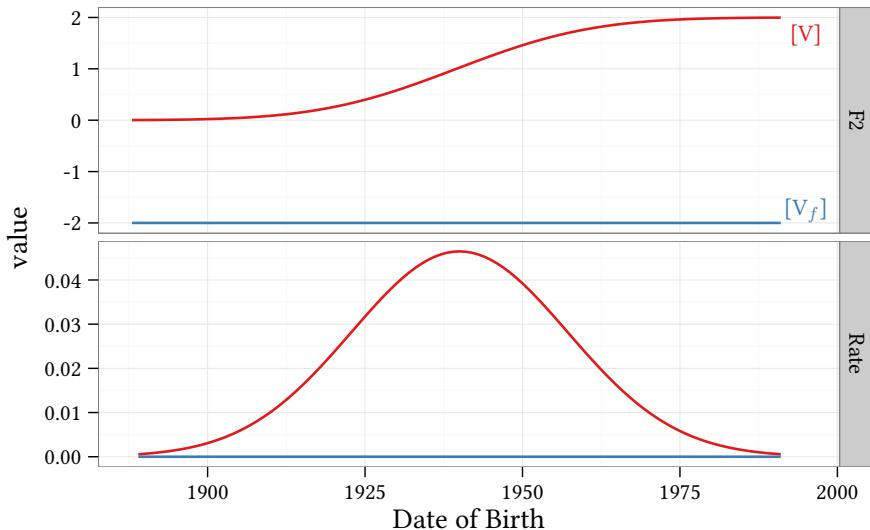


Figure 4.7: The rate of phonetic change in the context of phonological feature spreading.

By its very definition, the rate of change for $[V]$ reaches its maximum in the bottom facet of Figure 4.7 at the midpoint of the S-shaped curve in the top facet, because it is at the midpoint of the S-shaped curve that the change is progressing at its fastest. The rate of change for $[V_f]$ remains at 0 throughout, because it is not undergoing any phonetic change at all.

Another way to think about the relationship between the rate of change in the bottom facet of Figure 4.7 and the trajectory of change in the top facet is that the trajectory in the top facet represents the cumulative sum of values in the bottom facet. For example, the rate of change for

[V] in 1950 is approximately 0.04. This means that the predicted value of F2 in 1950 is equal to the value of F2 in 1949 plus 0.04. The value of F2 in 1949 was 1.42, so the predicted value of F2 in 1950 is $1.42 + 0.04 = 1.46$. To figure out how different the F2 of [V] is in 1950 from 1888 (the earliest point in time in these figures), we merely need to sum up all of the rates of change from 1888 to 1950, and add that to the value of F2 in 1888. In 1888, F2 was 0, and the sum of all by-year rates of change between 1888 and 1950 is 1.46, so the predicted value of F2 in 1950 is, again, $0 + 1.46 = 1.46$. Meanwhile, the rate of change for [V_f] in 1950 is 0, meaning the predicted value of F2 for [V_f] in 1950 is the value of F2 in 1949 plus 0; $-2 + 0 = -2$. The sum of all by-year rates of change from 1888 to 1950 for [V_f] is also 0, meaning that [V_f] is expected to have the same F2 in 1950 as in 1888.

A more technically accurate description of the relationship between the rate of change and the trajectory of change is that the rate of change is the first derivative of the trajectory of change. I will continue to describe the rate of change in terms of year-to-year differences for the sake of interpretability. However, keeping in mind that I am really trying to model $f'(x)$, where $f(x)$ is the trajectory of change, could be useful for technical advancements of these methods in the future.

The key takeaway from Figure 4.7 is that [V] and [V_f] have different rates of change, and as I argued in §4.1.1, this is only possible because they are phonologically distinct objects, and thus have different targets of phonetic implementation.

Figure 4.8 illustrates the expected dynamics of phonetic change if the contextual variants of /V/ were due to phonetic coarticulation. The solid line represents the movement of the target for [V_x] along F2. The arrows indicate the coarticulatory effect, shifting the productions of [V_y] back along F2 from the target for [V_x]. This coarticulatory effect remains constant over time, producing a trajectory for [V_y] which is yoked to [V_x], thus parallel to it over time.

The rates of change of two parallel trajectories, even if these trajectories are displaced upwards or downwards, will always be the same. This is represented in the bottom facet of Figure 4.8. At all points in time, [V_x] and [V_y] have the same rate of change because they are moving in parallel, because [V_y] is yoked to [V_x] because they share a target for V.

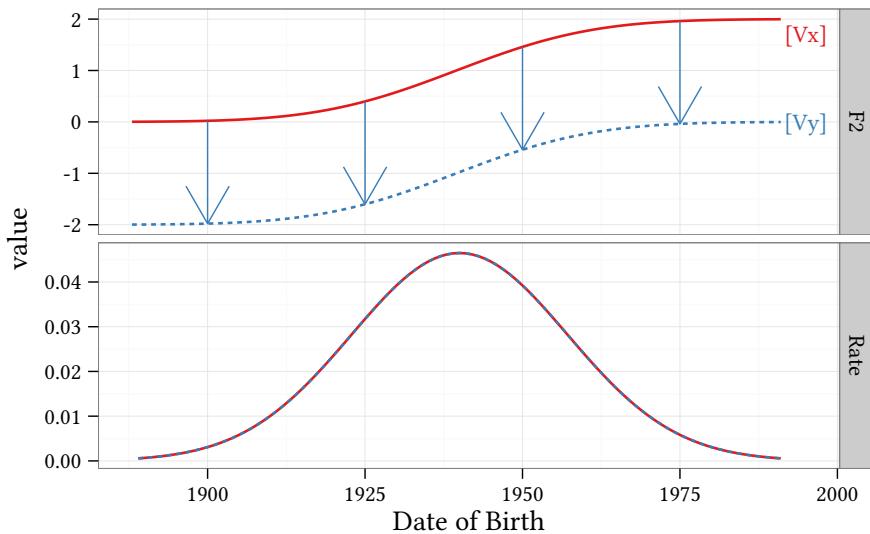


Figure 4.8: The rate of phonetic change in the context of phonetic coarticulation

The difference between phonological differentiation and phonetic coarticulation is large and qualitative. What I hope to have illustrated so far is that this qualitative difference can be connected to quantitative differences in the way the system changes over time. Specifically, for any given vowel which has two contextual variants, if we can estimate the rate of change of these two variants over time and determine whether they have a shared or different rate of change, then we can then use this information as an indicator of a qualitative difference.

Perhaps most importantly, we can utilize the comparison of rates of change to identify cases where phonetic coarticulation has been reanalyzed as phonological differentiation. That is, for some changes, the difference between [Vx] and [Vy] could have been originally due to phonetic coarticulation, but then speakers reanalyzed this difference as actually being due to a phonological process, with featurally distinct objects, [V] and [V_f], and targets. This process of reanalysis has been called “phonologization” (Hyman, 1976) or “stabilization” (Bermúdez-Otero, 2007), and is argued by some to be the primary source of naturalness in phonology (e.g. Cohn, 2006, 2007).

The effect this reanalysis would have on the dynamics of sound change is illustrated in Figure 4.9. At the beginning of the sound change, the difference in contextual variants of V is due to

phonetic coarticulation, and the trajectory of $[Vy]$ is yoked to $[Vx]$, causing them to have the same rate of change. The dark vertical line represents represents the time point when the coarticulatory effect is reanalyzed as being phonological. A process like (4.6) enters the phonological grammar, producing featurally distinct allophones, $[V]$ and $[V_f]$, which have independent targets of phonetic implementation. In this illustration, the trajectory of $[V]$ continues along its previous path, but $[V_f]$ ceases to undergo change.

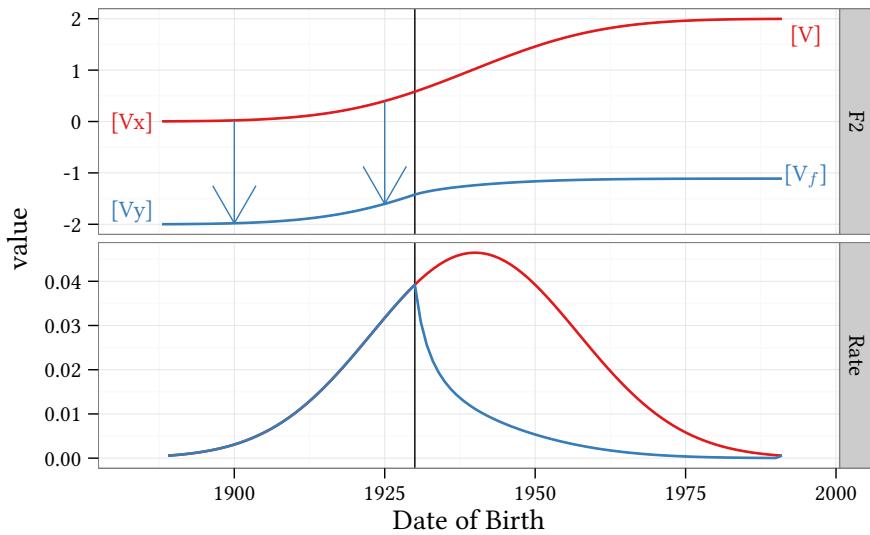


Figure 4.9: The reanalysis of phonetic coarticulation as phonological feature spreading, and its effect on the rate of phonetic change.

Looking at the trajectories alone, it would be difficult to pinpoint with much accuracy when the reanalysis occurred if were not indicated on the graph. The rates of change, on the other hand, indicate rather unambiguously a sharp point at which $[V_f]$ diverged from $[V]$. It is possible to model the trajectories directly using, for example, cubic regression splines, and comparing models where the trajectories are constrained to be the same to models where they are allowed to be different. This sort of modeling approach would tell us that in cases phonological feature spreading, like Figure 4.7, the trajectories differ significantly, while in the case of phonetic coarticulation, like Figure 4.8, they don't. However, this approach would also tell us that the trajectories differ significantly in cases where phonetic coarticulation has become reanalyzed as phonological feature spreading, like Figure 4.9. Given that we want to be able to disambiguate instances of all

three kinds of influences on sound change, and that in the case of reanalysis, we want to be able estimate a time point in the sound change when reanalysis occurred, a more complex approach is necessary, which involves directly modeling the rate of change.

I hope to have made clear, in this section, the possible diagnostic capacity of the rate of change. In principle, we should be able to not only identify qualitative differences through quantitative measure (i.e. the difference between phonetic coarticulation and phonological differentiation), but also identify the point in time where new qualitative options enter the grammar (i.e. the reanalysis of phonetic coarticulation as phonological differentiation).

4.3 The Model and the Data

This section will be devoted to the specifics of implementing a statistical model to estimate and compare the rates of change of different contextual variants, as well as the data behind the case studies I will be applying the model to.

4.3.1 The Model

As I stated above, we can conceptualize the rate of change as actually representing year-to-year differences along any particular phonetic dimension. Let's represent the rate of change for year l for a vowel in context k as δ_{lk} , which will be equal to the difference along the phonetic dimension between year $l - 1$ and l . This is the parameter of primary interest, specifically for particular years whether δ_{lk} is the same for different contexts. Contexts will be indexed by different values for k . The context $k = 1$ will always be some reference level context. For example, the first case study will focus on the effect of following nasals on /aw/. In this case, /aw/ followed by oral segments will be given index $k = 1$, and vowels followed by nasal segments will be given the index $k = 2$. Once we have estimated $\delta_{lk=1}$ and $\delta_{lk=2}$ for all l dates of birth, we will make the a quantitative comparison to see if $\delta_{lk=1} = \delta_{lk=2}$ or $\delta_{lk=1} \neq \delta_{lk=2}$. More precisely, we will be looking at the difference, $\delta_{lk=1} - \delta_{lk=2}$. There are three possible results for this comparison.

$$(4.7) \quad \delta_{lk=1} - \delta_{lk=2} > 0$$

This means that $\delta_{lk=1} > \delta_{lk=2}$ therefore $\delta_{lk=1} \neq \delta_{lk=2}$, therefore the vowel has different rates of change between contexts $k = 1$ and $k = 2$.

$$(4.8) \quad \delta_{lk=1} - \delta_{lk=2} < 0$$

This means that $\delta_{lk=1} < \delta_{lk=2}$ therefore $\delta_{lk=1} \neq \delta_{lk=2}$, therefore the vowel has different rates of change between contexts $k = 1$ and $k = 2$.

$$(4.9) \quad \delta_{lk=1} - \delta_{lk=2} = 0$$

This means that $\delta_{lk=1} = \delta_{lk=2}$, therefore the vowel has the same rate of change in contexts $k = 1$ and $k = 2$.

Now, δ_{lk} is not a directly observable variable in the data. Rather, it is a latent variable that we will be attempting to estimate from the data. For this reason, along with all of the usual constraints on statistical inference from a sample to a population, we will not be estimating precise values for $\delta_{lk=1} - \delta_{lk=2}$. Instead, we will estimating *credible intervals* for the value $\delta_{lk=1} - \delta_{lk=2}$. If the credible interval excludes 0, then our inference will be that it is more likely than not⁵ that $\delta_{lk=1} - \delta_{lk=2} \neq 0$. On the other hand, if the credible interval includes 0, our inference should be more cautious. It may actually be the case that $\delta_{lk=1} - \delta_{lk=2} \approx 0$, or it may be the case that the data is too sparse for either $k = 1$ or $k = 2$ to reliably determine otherwise.

As illustrated in Figures 4.7, 4.8 and 4.9, δ_{lk} should be modeled as a function of date of birth. However, I have no theoretically driven hypothesis about what the shape of that function ought to be. As such, I made the decision to model δ_{lk} using b-splines. I chose b-splines over other kinds of curve fitting because they are relatively easy to implement, conceptually easy to understand, and flexible in the kinds of curves they can approximate. Fitting a curve with b-splines begins by defining the “*basis*” of the curve. In the context of curve fitting, “*basis*” has a technical meaning of approximately a collection of curves which are then scaled and summed over to produce the final curve. Figure 4.10 displays the b-spline basis used in all of the models in this chapter, which was constructed with the `splines` package in R (R Core Team, 2012). This particular basis consists of three cubic polynomial curves which are evenly spaced along the time dimension, and one linear intercept term.

⁵In fact, for the credible intervals displayed in this work, it will be 95% more likely than not.

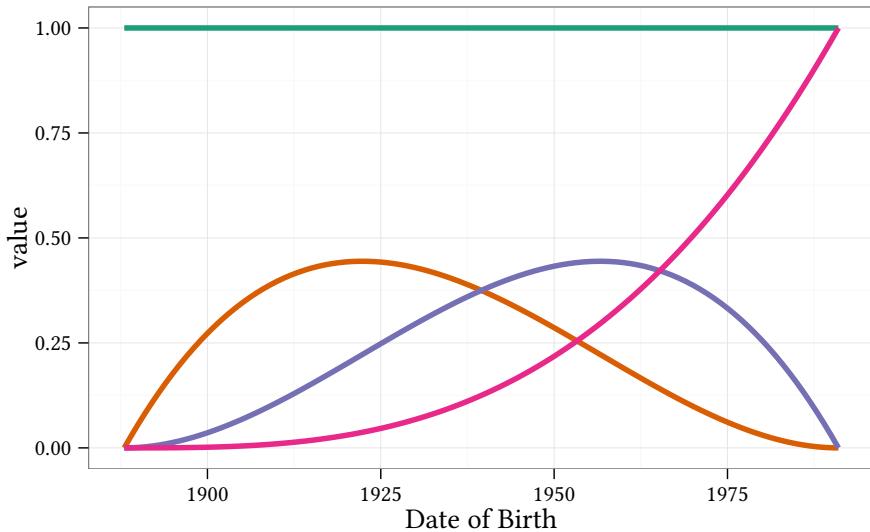


Figure 4.10: B-spline basis used in all following models.

After establishing the basis of the b-spline, you then estimate *weighting coefficients* for each curve in the basis. Usually, the weighting coefficients will be estimated from the data, but in this illustration, 4 coefficients were randomly chosen from a normal distribution. You then scale each polynomial by multiplying it by its corresponding weighting coefficient. The weighted form of the basis is represented in the top facet of Figure 4.11. In the final step, you sum across the polynomial along the x-axis, resulting in the final b-spline fit, which is represented in the bottom facet of Figure 4.11. Figure 4.12 displays five more b-spline fits based on more randomly generated weighting coefficients in order to provide a qualitative sense of how smooth b-spline fits with the basis in Figure 4.10 will be.

The degree of wigginess of a b-spline fit is highly dependent on the size of the basis. For example, Figure 4.13 displays the kind of curve that a larger b-spline basis could fit. I will be restricting my modeling of δ_{lk} to the smaller basis displayed in Figure 4.10 for the following reasons.

- (4.10) As the size of the basis increases, the number of weighting coefficients increases, and the over all uncertainty about the final fit of the curve increases.

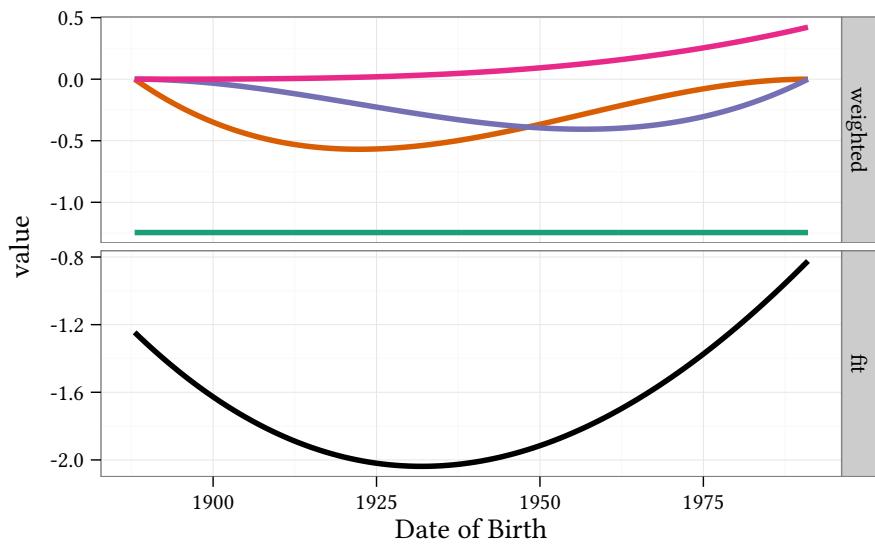


Figure 4.11: Weighted b-spline basis, and resulting spline fit.

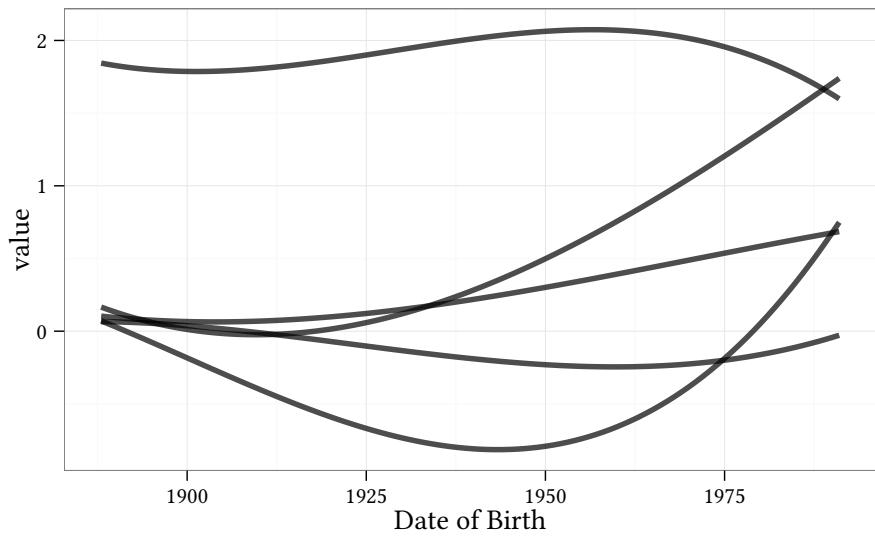


Figure 4.12: Five randomly generated b-spline curves

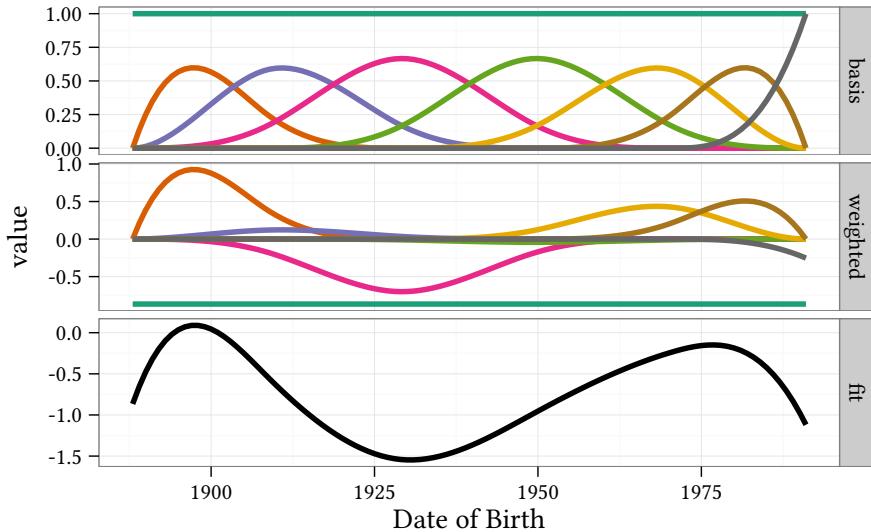


Figure 4.13: The effect of a larger basis on the fit’s ‘wiggleness’.

(4.11) Since δ_{lk} is a latent variable, there is already a higher degree of uncertainty built into its

estimation.

(4.12) Additionally, since δ_{lk} represents the first derivative of the trajectory of change, it can

afford to be relatively simpler than the actual trajectory, since $f'(x)$ is always one degree less than $f(x)$.

I will represent the fact that δ_{lk} is modeled by a b-spline smooth of date of birth, which is also the index for l , as follows.

$$\delta_{lk} = b.\text{spline}(l) \quad (4.13)$$

After estimating δ_{lk} for every date of birth, we then need to estimate the expected value along the phonetic dimension for that date of birth. That is, if the change we are modeling is /ow/ fronting along F2, δ_{lk} will represent how far /ow/ fronted along F2 between the years $l - 1$ to l , but we also need to estimate what the actual value of F2 is in year l . As was discussed in §4.2, this can be done by taking the cumulative sum of δ_{lk} from 1888 up to year l , then adding it to

the value of F2 in 1888. The cumulative sum will be represented by Δ_{lk} , the value in 1888 will be represented by β_k , and the expected value in year l will be represented by μ_{lk} .

$$\Delta_{lk} = \sum_{x=1888}^l \delta_{xk} \quad (4.14)$$

$$\mu_{lk} = \beta_k + \Delta_{lk} \quad (4.15)$$

At this point, μ_{lk} represents the expected phonetic target for a vowel in context k for a speaker born in year l . However, it would not be expected for *all* speakers born in year l to have the precise target of μ_{lk} . Obviously, inter-speaker variability exists for all manner of systematic reasons, some of which could be incorporated into the model, like socio-economic class, education, etc. Just as obviously, there are systematic causes of inter-speaker variation that we cannot include in the model because it didn't occur to us to document them, we have yet to operationalize measures for them, or they are in some sense immeasurable, related to the accidental personal history of every individual. Finally, even with a full accounting of all possible factors that predict inter-speaker variation, and well formulated operationalizations and measurements of those factors, there will always be some variation between individuals which is irreducible.

For these reasons, we will add an additional layer to the model, where we estimate phonetic targets for every individual speaker in the corpus, which will be represented as μ_{jk}^s , where j is an index for each speaker. These speaker-level parameters will be drawn from a normal distribution centered around μ_{lk} . The variance of the distribution will be another parameter in the model σ_k . The reason we want to include σ_k as a parameter in the model is that we want to allow speakers to be as similar to each other, or as different from each other as is warranted by the data. Notice that σ_k is also indexed by the context k . This means that inter-speaker variation can be greater or lesser for each context under question. In the following equations, DOB_j represents the date of birth for speaker j .

$$DOB_{1,2,\dots,n.speaker} \quad (4.16)$$

$$l = DOB_j \quad (4.17)$$

$$\mu_{jk}^s \sim \mathcal{N}(\mu_{lk}, \sigma_k) \quad (4.18)$$

Additionally, we should recognize that speakers will differ in the degree to which the individual tokens they produce are scattered around their target. Some speakers may have very small variance, with most of their productions being clustered tightly around their basic target, μ_{kj}^s , while other speakers may have much larger variance. As such, we will also be estimating speaker-level variances, which will be represented as σ_j^s .

An additional point of complexity to the data is that not only is it generated by many different speakers, but also represent many different lexical items. Whether or not lexical items play an important role in sound change over and above environmental conditioning is a long, and ongoing debate (Labov, 1981, 1994, 2010a; Pierrehumbert, 2002; Bybee, 2002 *inter alia*). Regardless of whether or not lexical items can have individualized phonetic targets, I will be including by-word random effects in this model for much the same reason as why by-speaker random effects were included. It is certainly the case that there are systematic properties of lexical items which affect their phonetic realizations which we have not accounted for, and are therefore missing from the model. Therefore, we will be estimating by-word random effects drawn from a normal distribution centered around 0, with a variance parameter which will be estimated on the basis of the data. The random effect for each word will be indexed by m , and will be represented as μ_m^w . As can be seen in the equation below, μ_m^w is not sensitive to any properties of the speaker, including date of birth, making it time insensitive. It would be ideal to model the effect of a word as being variable over time, to see if it changes or remains stable, but the model as I've laid it out up to this point is already very complex, and making the by-word effects time sensitive would *minimally* involve adding two parameters to the model for every lexical item: slope and intercept. Therefore, I will be backing off from an ideal model of lexical effects to a merely sufficient one.

$$\mu_m^w \sim \mathcal{N}(0, \sigma^w) \quad (4.19)$$

Finally, we come to the raw data layer of the model. The raw acoustic data will be represented by y_i , where i is an index for every observation. The total number of observations is represented

as n , J is a vector of speaker indices, K is a vector of context indices, and W is a vector of word indices. We will be adding speakers' expected phonetic target for a vowel in context k , represented by μ_{jk}^s , to the word level effect, μ_m^w , to arrive at the expected target for observation y_i . Of course, any particular observation from a particular speaker of a particular word will not precisely be equal to $\mu_{jk}^s + \mu_m^w$ for all of the reasons which have already been stated, so we will actually be presuming that y_i is drawn from a normal distribution centered around $\mu_{jk}^s + \mu_m^w$ with a speaker specific variance, σ_j^s , which was mentioned above.

$$y_{1,2,\dots,n} \quad (4.20)$$

$$J_{1,2,\dots,n} \quad (4.21)$$

$$K_{1,2,\dots,n} \quad (4.22)$$

$$W_{1,2,\dots,n} \quad (4.23)$$

$$j = J_i \quad (4.24)$$

$$k = K_i \quad (4.25)$$

$$m = W_i \quad (4.26)$$

$$y_i \sim \mathcal{N}(\mu_{jk}^s + \mu_m^w, \sigma_j^s) \quad (4.27)$$

Human Readable Form

This model of the rate of change has three levels. At the highest level, the year-over-year differences are estimated using non-linear curve fitting. I didn't assume that the rate of change was constant across the lifespan of the phonetic change because, in fact, all three of the changes I look at in this chapter move in one direction, stop, then reverse, and also, the relative timing of when contextual variants diverge in their rate of change is of key interest. At the next level, the estimated phonetic targets of each speaker are estimated. The expected phonetic target for a speaker born in a particular year is estimated by summing up the year-over-year differences from the first layer of the model. The phonetic targets of the actual speakers in the model are assumed to be

normally distributed around the expected target for their date of birth. By-word random errors are also assumed to be normally distributed around 0. The third, and lowest level, treats each individual measurement as being drawn from a normal distribution centered around the specific speaker's phonetic target plus the particular word's random error.

Some readers may be more familiar with the syntax of mixed effects linear models as implemented in the `lme4` R library. Faux-lme4 syntax for this rate of change model is provided in (4.28) and (4.29). It includes random intercepts for speaker and word.

(4.28) `rate_of_change ~ b_spline(DOB)`

(4.29) `F2 ~ sum(rate_of_change) + (1|Speaker) + (1|Word)`

4.3.2 Implementing the model

The model I have just described does not easily submit to a reformulation as a linear regression, or even in terms of modeling techniques like generalized additive models. As such, I have implemented it in Stan (Stan Development Team, 2012). Stan is a package designed to implement graphical Bayesian models with Hamiltonian Monte Carlo (Hoffman and Gelman, 2011). Providing a precise description of HMC is well beyond the scope of this dissertation. Generally speaking, HMC is closely related to Markov chain Monte Carlo methods of model estimation, for which Kruschke (2011) is an excellent introduction. In an iterative process, the system samples possible values for the parameters it's trying to estimate from a probability distribution which is in part determined by its prior probability, the probability of the other parameter values estimated so far, and the observed data. After a sufficient number of iterations, the samples produced by the system will approximate the posterior probability distribution of the parameters, which is what we will use for our inferences. As it is an iterative process, we want to be sure that it is not sensitive to its initial values, so the model will be fit multiple times with different random initializations, and the results compared across the fits (or chains) to verify that they have converged on the same values. Figure 4.14 illustrates the convergence of three chains to stable distribution. The parameter being estimated in this case is δ_{lk} (the rate of change in year l for context k) for women in 1940 for /aw/ in pre-oral contexts. The top facet represents the full trace of the three chains. As can be seen,

in the first few iterations, the values being estimated for δ_{lk} vary broadly, but rapidly converge to a narrower range. Not all parameters converge this quickly, so as a general practice, the first half of the samples are discarded as a “burn-in.” The second half of the samples are taken to be representative of the posterior distribution, which is represented in the bottom facet of Figure 4.14.

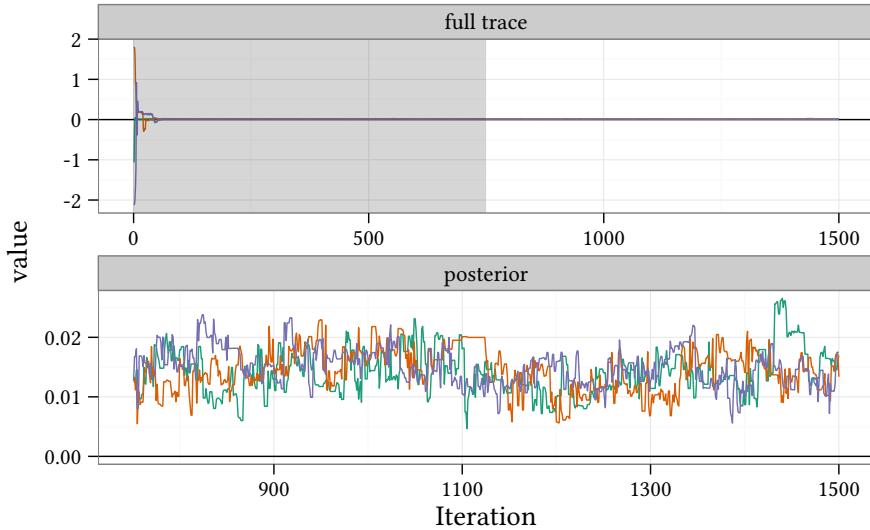


Figure 4.14: The full trace for three chains estimating δ_{lk} for $l = 1940$, and the sample approximating the posterior.

There are a number of different diagnostics for determining how well converged a model is. Note, these are not diagnostics of how well the model estimates match reality, which is unknown, but rather, how *consistent* the model’s estimates are. I will be employing the Gelman and Rubin Potential Scale Reduction Factor, represented by \hat{R} , which compares the between-chain variance to the within-chain variance. Values of \hat{R} close to 1 indicate good convergence, and the example in Figure 4.14 has $\hat{R} = 1.06$.

As a Bayesian model, it’s necessary to define prior probability distributions over the parameters it is going to estimate. I’ve already specified some of the model priors above. For example, I specified that the speaker-level target estimates, μ_{jk}^s should be normally distributed around the community-level estimate for that speaker’s date of birth, μ_{lk} . However, there are many parameters for which I have not mentioned what their prior probability distribution should be, like

the variance parameters, σ , σ_j^s , σ^w , or the b-spline weighting coefficients. These parameters, and any others not explicitly mentioned in the description above, were given non-informative, or weakly informative priors. Specifically, scale and variance parameters were given a uniform prior between 0 and 100, $\sim U(0, 100)$, and all other parameters were given a normal prior with mean 0 and standard deviation 1,000, $\sim \mathcal{N}(0, 1000)$. Given the scale of the data, which is z-score normalized Hz measurements, these constitute, at most, weakly informative priors.

4.3.3 I have just described a generative model

The model I have described is called a generative model in statistical terminology, because it describes a model of how the observed data was generated. That is, it models observations as being drawn from speakers speaking specific words, and speakers as being drawn from a larger and dynamically changing population. However, I believe there is also a felicitous convergence of terminology here with “generative” as it is used in Linguistics. To begin with, the specification of any statistical model is theory laden, and the reason I have specified the model above is driven primarily by the linguistic theory I want to evaluate, which is based on generative phonology and phonetics. Moreover, some of the parameters in the model correspond nicely to theoretical concepts in linguistics. Specifically, μ_{lk} , which represents the expected phonetic target of a speaker born in year l , could be understood as representing the “community grammar” in the sense of Weinreich et al. (1968). Alternatively, it could just as easily be conceived of as representing the knowledge of

...an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. (Chomsky, 1965)

In the model, speaker-level factors such as memory limitations, distractions etc. are factored out by σ_{jk}^s to arrive at the idealized knowledge of each speaker, μ_{jk}^s . Community level factors, such as unaccounted for heterogeneity, is factored out by σ_k , to arrive at the idealized knowledge of an idealized speaker, μ_{lk} . The goal of this model is to determine what factors can account for

the idealized knowledge of an idealized speaker, which is also a goal of generative linguistics as I understand it.

4.4 Case Studies

All of the the cases studies presented here are based on data drawn from the Philadelphia Neighborhood Corpus (Labov and Rosenfelder, 2011). The measurements used are those produced by the FAVE suite (Rosenfelder et al., 2011), but additional contextual information has been collected from the PNC raw data.

Figure 4.15 is presented as background, and represents the trajectories of sound change in the 1970s as determined by the LCV study in Philadelphia. I will be examining the conditioning factors on /aw/, /ow/ and /uw/ here. Table 4.1 provides a broad IPA transcription for these vowel classes, their corresponding Wells Lexical Set labels, and an approximate transcriptions defining the range of phonetic variation.

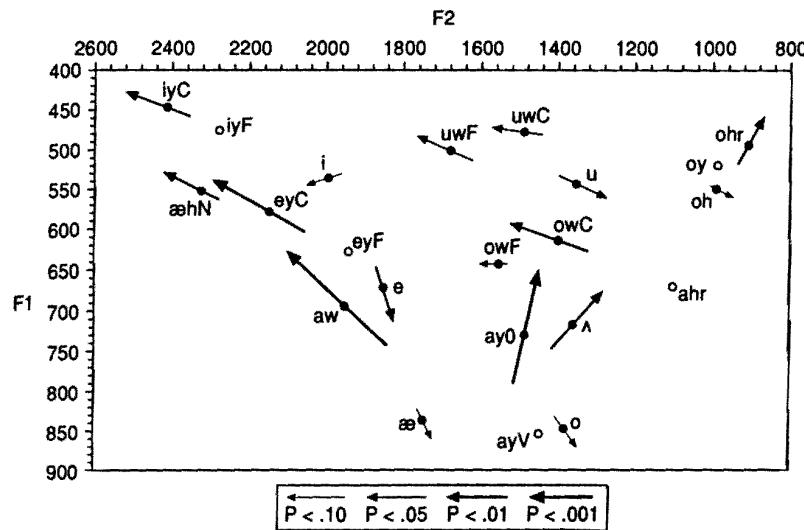


Figure 4.15: The Philadelphia Vowel System in the 1970s. From Labov (2001).

Before moving forward, I should note that for all of the following analyses, vowel tokens that were either word initial, or co-extensive with the word were excluded. I did this, in part, to exclude as many cases which could be attributable to errors in forced alignment, but also to reduce

Label	Broad IPA	Wells Lexical Set	Range of Variation
/aw/	[æw]	MOUTH	[eɔ]~[æʊ]
/ow/	[ow]	GOAT	[əʊ]~[o:]
/uw/	[uw]	GOOSE	[iʊ]~[u:]

Table 4.1: The case studies in this chapter, a broad IPA transcription, Wells Lexical Set labels, and IPA transcription of the range of variation.

the number of cases being examined. As we will see, the vowel in each case study is already subdivided in many ways, and including parameters for word initial and co-extensive tokens would have expanded the size of the statistical models further, without any clear advantages in return.

4.4.1 /aw/

In the 1970s, the fronting and raising of /aw/ along the front diagonal of the vowel space was identified as a vigorous change in progress (Labov, 2001). In the PNC, /aw/ raising has been found to be reversing, starting with speakers born around the 1960s. It also exhibits strong sociolinguistic conditioning, with a large difference between women, who are more advanced, and men (Labov et al., 2013). Figure 4.16 plots the basic, smoothed trajectory for /aw/ in $F1 \times F2$ space, overlaid on the full vowel triangle for context.

The PNC group has found that the best way to capture movements along the front diagonal of the vowel space is with a diagonal measure given as $(F2 - \beta F1)$, where β depends on the transformed scale of the data. In the z-score normalized space, which I am presenting here, the optimal value for β is 1, so Diag, in all future figures and statistics, is simply $(F2 - F1)$. Figure 4.17 displays the basic trajectory of /aw/ over time. Each point represents the mean value of Diag for one speaker.

The largest conditioning factor on the raising and fronting of /aw/ is whether it is followed by a nasal or oral segment (Labov et al., 1986). Figure 4.18 displays the mean values for speakers for /aw/ in pre-oral and pre-nasal contexts. Throughout the entire change, [awN] is consider-

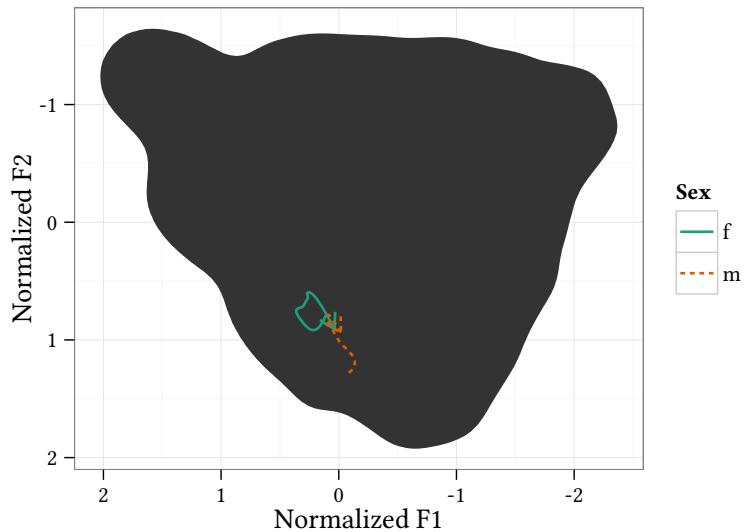


Figure 4.16: /aw/ Trajectory in $F_1 \times F_2$ Space

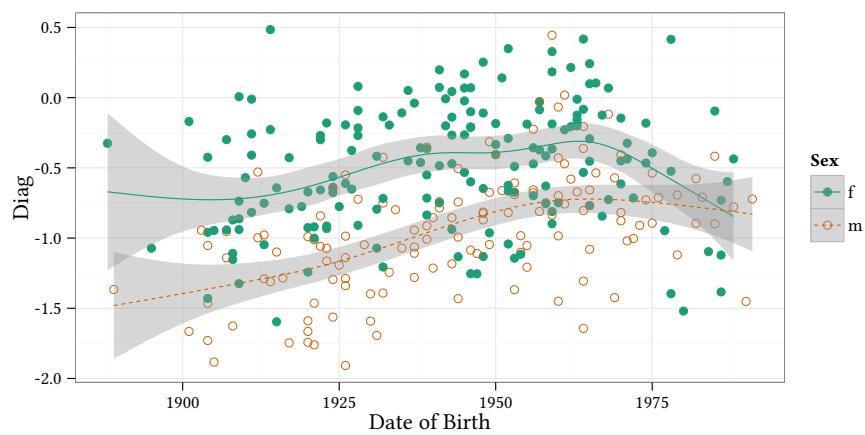


Figure 4.17: /aw/ change trajectory

ably more advanced along the Diag scale for both men and women. At this point, it appears impressionistically that the effect of following nasals remains consistent across the 20th century.

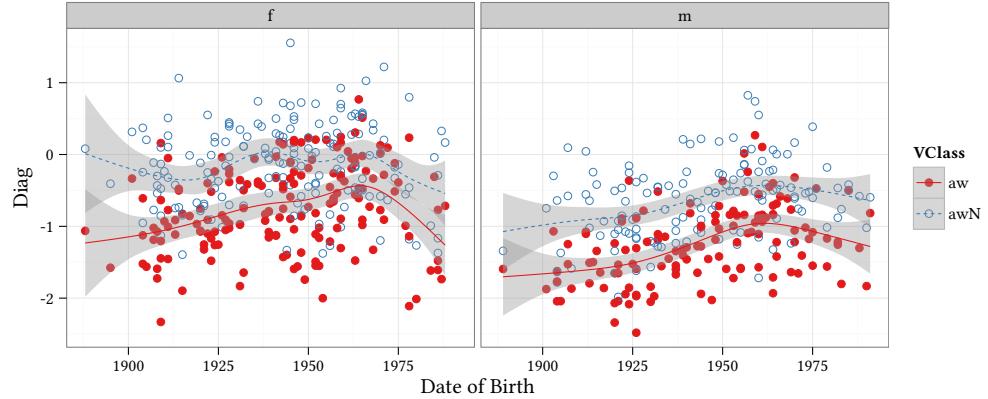


Figure 4.18: The effect of following nasals on /aw/.

Given that following nasals have such a strong effect on /aw/, I also coded /aw/ according to whether it was preceded by a nasal (to be represented by [Naw]), and whether it was sandwiched by nasals (to be represented by [NawN]). Also, for maximal parallelism between the rest of the analyses in this section, I also coded /aw/ for whether or not it was word-final, or followed by an /l/. Table 4.2 describes the coding criteria.

Variant	Criteria
awN	Is followed by /m, n, ɳ/ within the same word
Naw	Is preceded by /m, n, ɳ/, and is not word final
NawN	Is preceded and followed by /m, n, ɳ/ within the same word
awF	Is word final, and not preceded by /m, n, ɳ/
NawF	Is word final, and preceded by /m, n, ɳ/
awL	Is followed by /l/ within the same word
aw	Remaining cases

Table 4.2: Coding criteria for /aw/

Table 4.3 displays the token counts of each variant in the corpus. Given that [awL], [Naw] and [NawN] are relatively rare, and that the parameter δ_{lk} is already fairly abstracted away from the data, I will be excluding these contexts from further analysis in this section. That leaves [aw], [awN], [NawF] and [awF]. The results for [NawF] should be taken with some caution, however,

Variant	N	Word Types
aw	6382	150
awN	5494	147
NawF	2504	1
awF	1377	11
NawN	183	21
Naw	92	12
awL	14	5
total	16046	343

Table 4.3: Token counts of each variant

because even though it is relatively frequent, it consists entirely of tokens of the word *now*. As I mentioned in §4.3, a reference level must be chosen to compare the other variants to. In this case, that will be /aw/. As a first pass visualization of the data, Figure 4.19 plots cubic regression splines over speaker means for each of these /aw/ variants. For the most part, all variants seem to follow the same trajectory transposed up and down with the possible exception of [awF] for women. The extreme wobbliness of [awF] for women is almost certainly not “real”, but is rather a common result of sparse data for these curve fitting methods.

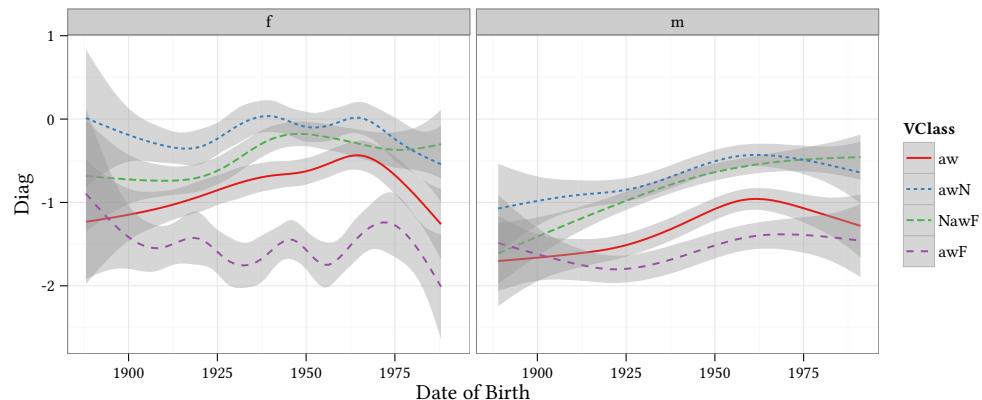


Figure 4.19: The /aw/ variants to be modeled.

Model Fit

As a first step to evaluating the quality of the model fit, we'll first examine the trajectory of change it predicted. Figure 4.20 displays the distribution of \hat{R} values for the estimated trajectories of [aw], [awN], [awF] and [NawF]. Values of \hat{R} indicate good convergence of the model. Unfortunately, the model appears not to have converged well at all for [NawF], with most of its \hat{R} values being greater than 2. As mentioned before, this is probably due to the fact that [NawF] is represented by just 1 lexical item in the corpus: *now*. For this reason, the data is actually sparser for [NawF] than the raw data might suggest, and the model estimation would face considerable ambiguity between attributing the target of [NawF] to its specific μ_{lk} value, or to its word-level effect, μ_m^w . Moving forward, [NawF] will be excluded from the analysis.

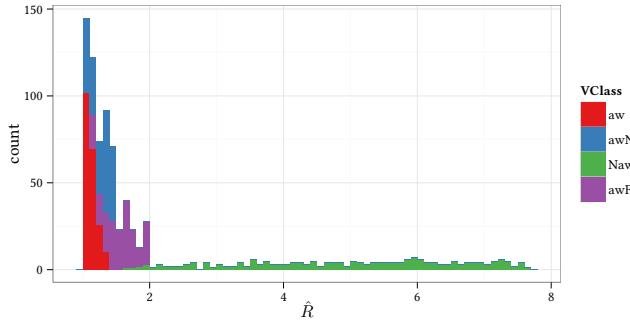


Figure 4.20: \hat{R} for the predicted trajectories of /aw/ variants.

Figure 4.21 displays the 95% Highest Posterior Density intervals for the predicted trajectories of change for [aw], [awN] and [awF]. There is a slightly larger probability range for the trajectory of [awF], but over all, these trajectories seem to fairly well fit, approximating closely the trajectories in Figure 4.19, suggesting that the model has not “blown up.”

Figure 4.22 plots the estimates of the parameter of central interest, δ_{lk} , representing the year-to-year differences. All three variants appear to share approximately the same rates of change, but [aw] appears to be slightly more exaggerated than the other two. In this figure, and the ones that follow, the color of the lines along the edges of the 95% HPD indicate whether or not 0 is excluded. As we can see in Figure 4.22, all three variants of /aw/ have significantly positive rates of change starting somewhere around the mid 1910's and continuing into the 1950's. For [awF]

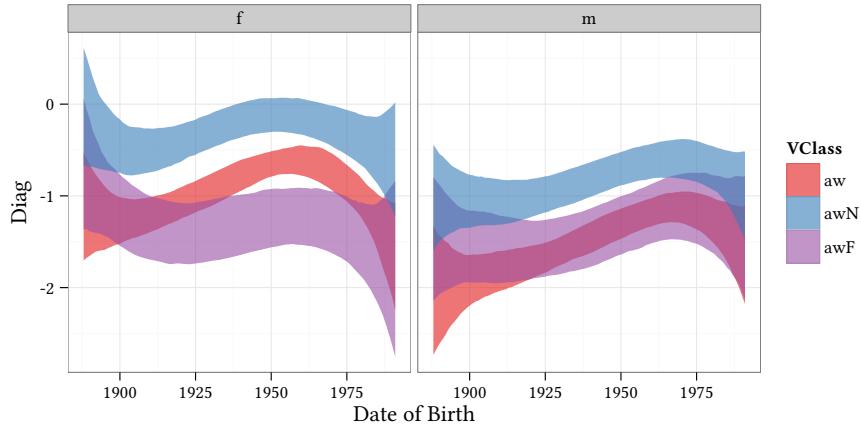


Figure 4.21: Predicted trajectories of change for /aw/ variants.

for women, the 95% HPD never actually excludes 0 in this time period, but its over-all trend is the same. For women, the rate of change for [aw] and [awN] turns negative as the change begins to reverse in the early 1960's.

However, the key comparison to make is whether [awF] and [awN] have reliably different rates of change from [aw]. This comparison is made in Figure 4.23, and as can be seen, neither [awF] nor [awN] exhibit considerable differences from [aw]. There is a brief period of about 5 years where [awF] seems to be changing more slowly than [aw], and in fact, looking at Figure 4.21, this is because [awF] appears to be moving downwards. Given the fact that this trend is so brief (less than 10 years), and that it is located so early in the sample, where there are fewer speakers, I'll attribute this blip to the the idiosyncrasies of a few speakers' data, rather than to a real trend.

Discussion

In the cases where there was enough data to make the comparison, it appears as if the different contextual variants of /aw/ share the same rate of change. Over the course of 100 years, [aw], [awN] and [awF] follow parallel trajectories, a remarkable fact in and of itself. My conclusion for /aw/ is that its most notable conditioning factor, the presence of a following nasal, is due to

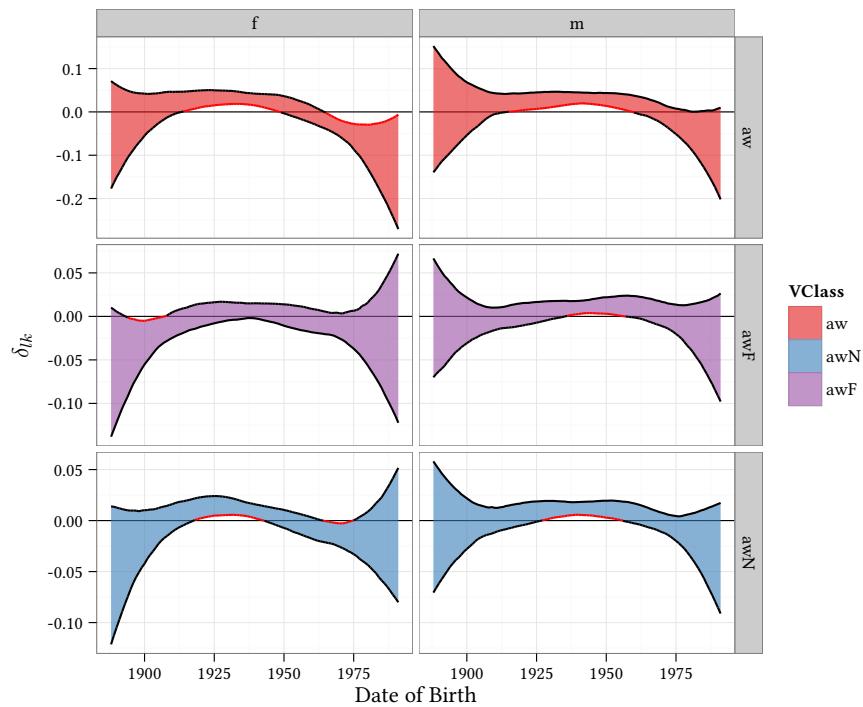


Figure 4.22: Year-to-year differences for variants of /aw/. Note: y-axis ranges differ across each horizontal set of facets.

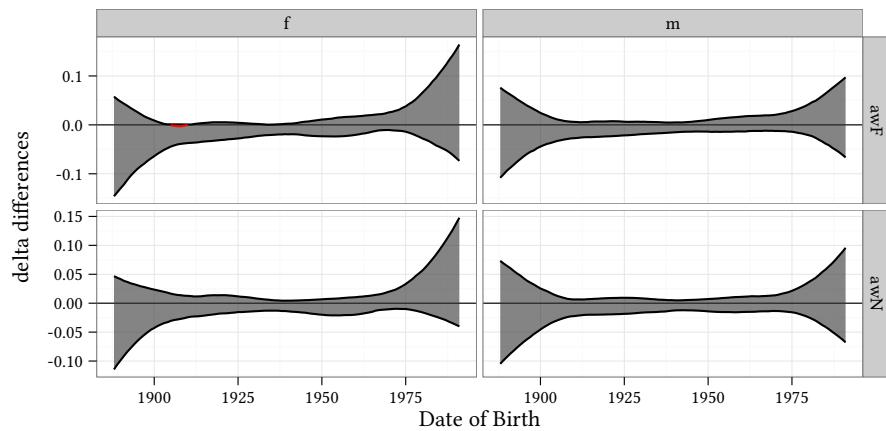


Figure 4.23: Rate of change differences from /aw/. Note: y-axis ranges differ across each horizontal set of facets.

phonetic coarticulation, not due to any categorical phonological process. That is, there is just one target for /aw/ that changes over time, and that target is merely shifted upwards in production by the presence of a following nasal.

4.4.2 /ow/

The next case study in this chapter is /ow/ fronting. Again, /ow/ fronting was found to be a change in progress in Philadelphia in the 1970's that has since begun reversing (Labov, 2001; Labov et al., 2013). Figure 4.24 plots the trajectory of /ow/ in the F1×F2 space. Women underwent a fronting change which has reversed, but the pattern for men is a bit more ambiguous.

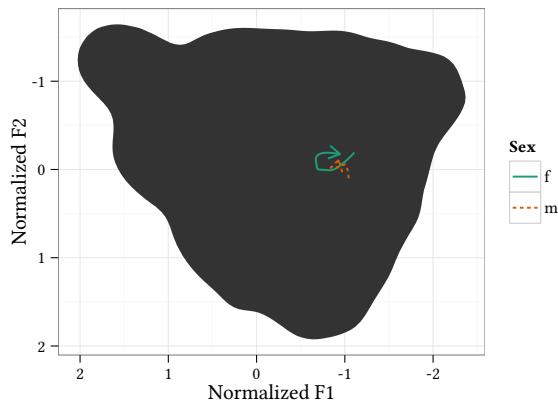


Figure 4.24: /ow/ Trajectory in F1×F2 Space

As a fronting change, the acoustic measure being used in this section is simply normalized F2. There are two major conditioning factors on /ow/ of note. First is whether or not /ow/ is absolute word final, which favors a more fronted form of /ow/, and whether or not it is followed by /l/, which favors a backer form of /ow/.

Figure 4.25 plots speaker means for these /ow/ variants. The first thing to note is that there is much stronger sociolinguistic differentiation between men and women for /ow/ than there was for /aw/. For /aw/, men lagged behind women, but were still participating in the change. For /ow/, it does not appear as if men undergo any change at all. The directions of the contextual effects are the same between men and women, but there is not much diachronic pattern to speak of for men. The second thing to note is that for women, the differentiation of [ow] and [owL] looks

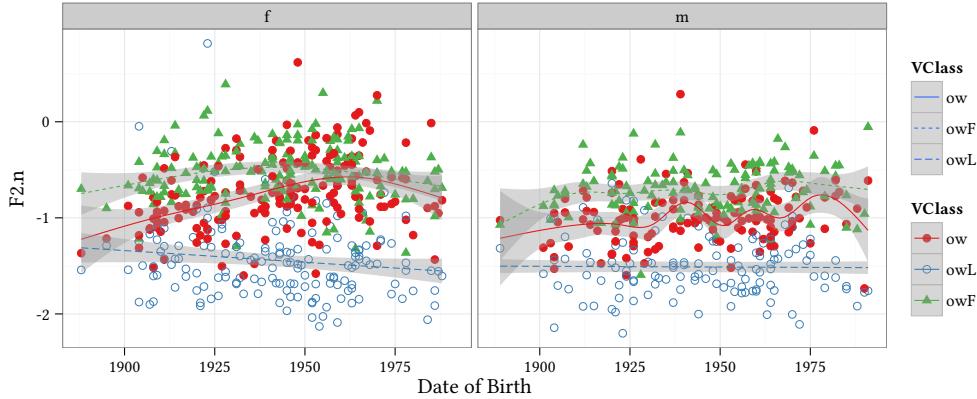


Figure 4.25: The effect of following /l/ and a word boundary on /ow/.

strikingly similar to the hypothetical patterns of phonological processes interacting with sound change from §4.2.

To keep the results for /ow/ as comparable to the results for /aw/ as possible, I coded /ow/ in the same way. Table 4.4 displays the coding scheme for /ow/, and Table 4.5 displays the total number of tokens for each variant.

Variant	Criteria
owN	Is followed by /m, n, ɳ/ within the same word
Now	Is preceded by /m, n, ɳ/, and is not word final
NowN	Is preceded and followed by /m, n, ɳ/ within the same word
owF	Is word final, and not preceded by /m, n, ɳ/
NowF	Is word final, and preceded by /m, n, ɳ/
owL	Is followed by /l/ within the same word
ow	Remaining cases

Table 4.4: Coding criteria for /ow/

As with /aw/, the /ow/ variant sandwiched by nasals is too low frequency to include in the model, and will be excluded from here on out. All of the other variants are relatively high frequency. Unsurprisingly, the highest frequency variant, [NowF], is dominated by the lexical items *no* and *know*, but is not exclusively constituted of them, so it should not exhibit the same ill fitting that [NawF] did. Figure 4.26 displays cubic regression splines over speaker means for the /ow/

Variant	N	Word Types
NowF	12956	33
owF	11092	200
owN	6091	164
ow	5629	592
owL	1946	153
Now	1836	123
NowN	95	11
total	39645	1253

Table 4.5: Token counts of each variant

variants to be included in the model. In the model, as with /aw/, the reference variant will be [ow], which corresponds to non-word final /ow/ that is not followed by /l/, and is neither preceded nor followed by a nasal consonant.

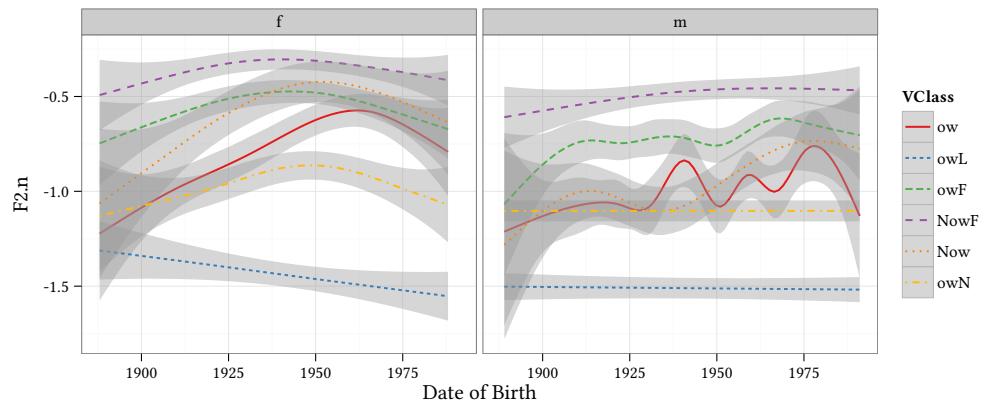


Figure 4.26: The /ow/ variants to be modeled.

Model Fit

Again, we'll examine how well the model estimated the basic trajectories of /ow/. Figure 4.27 displays the \hat{R} convergence estimates for the trajectories, broken down by variant. As with /NawF/, /NowF/ is the least well converged variant, but in this case, its \hat{R} values are not dire, so we will keep /NowF/ in for the rest of the analysis.

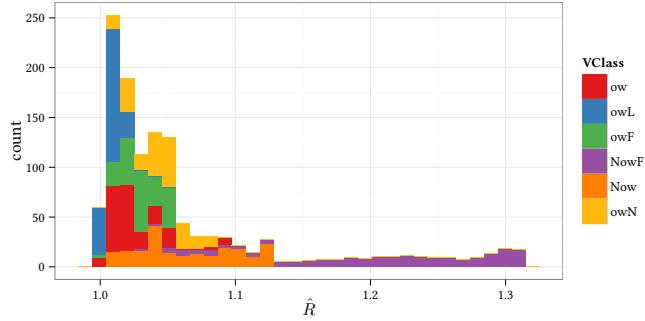


Figure 4.27: \hat{R} for the predicted trajectories of /ow/ variants.

Since there are six different variants of /ow/, I will not be plotting them over each other. Instead, Figure 4.28 plots a row of facets for each /ow/ variant and superimposes the predicted trajectory for [ow] on each one. All of the trajectories appear to be fit to comparable degrees of certainty, and look very similar to the trajectories in Figure 4.26. Again, there seems to be a striking difference between [owL] and all other variants. Every other variant of /ow/ (for women at least) appears to undergo some sort of fronting and subsequent backing, and [owL] is completely divergent from this pattern. The divergence of [owL] appears equally strongly when looking at the rates of change in Figure 4.29.

For women, [ow], [owF], [NowF], [Now] and [owN] all exhibit a very clear pattern of a positive rate of change in fronting beginning somewhere near the turn of the century, followed by a reversal starting around the 1960's. In contrast, the rate of change for [owL] is plausibly 0 throughout the entire century. The pattern for men is much more ambiguous. There is some hint of fronting for some variants centered approximately around the 1950's, but it is very subtle.

Finally, we come to the differences in the rates of change between [ow] and the other variants in Figure 4.30. Men have virtually no difference between [ow] and the other variants, so the rest of this discussion will focus exclusively on women.

As expected, [owL] has a reliably different rate of change from [ow] almost from the very beginning of /ow/ fronting. The earliest date of birth where [ow] has a rate of change reliably greater than 0 (as depicted in the top row of facets in Figure 4.29) is 1906, and the earliest [owL]

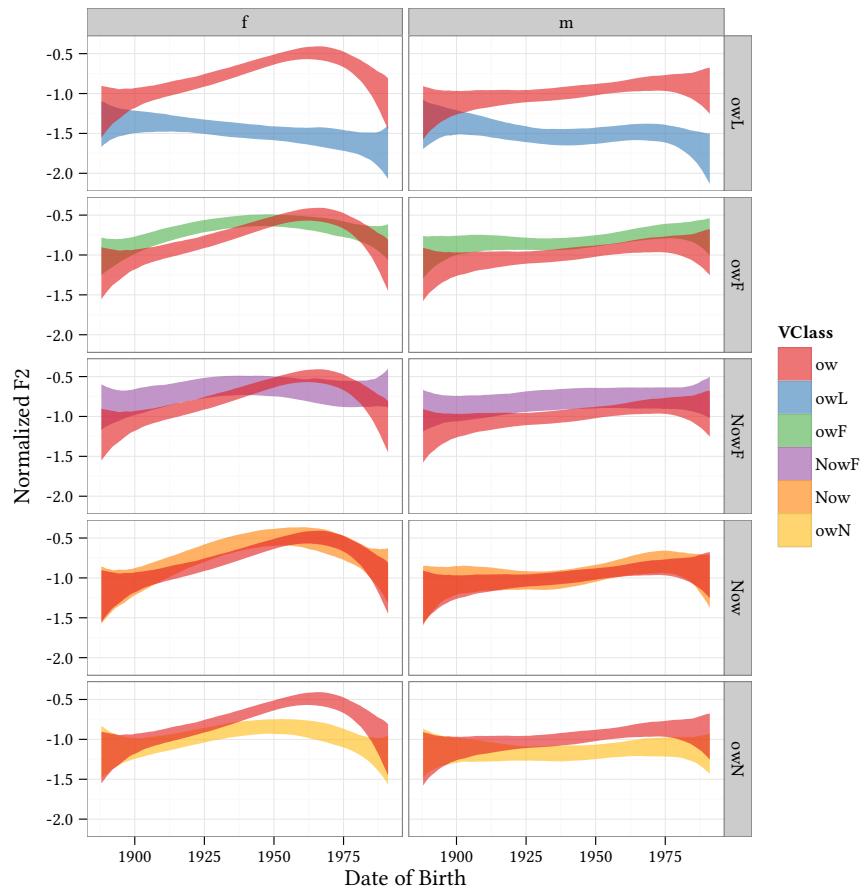


Figure 4.28: Predicted trajectories of change for /ow/ variants.

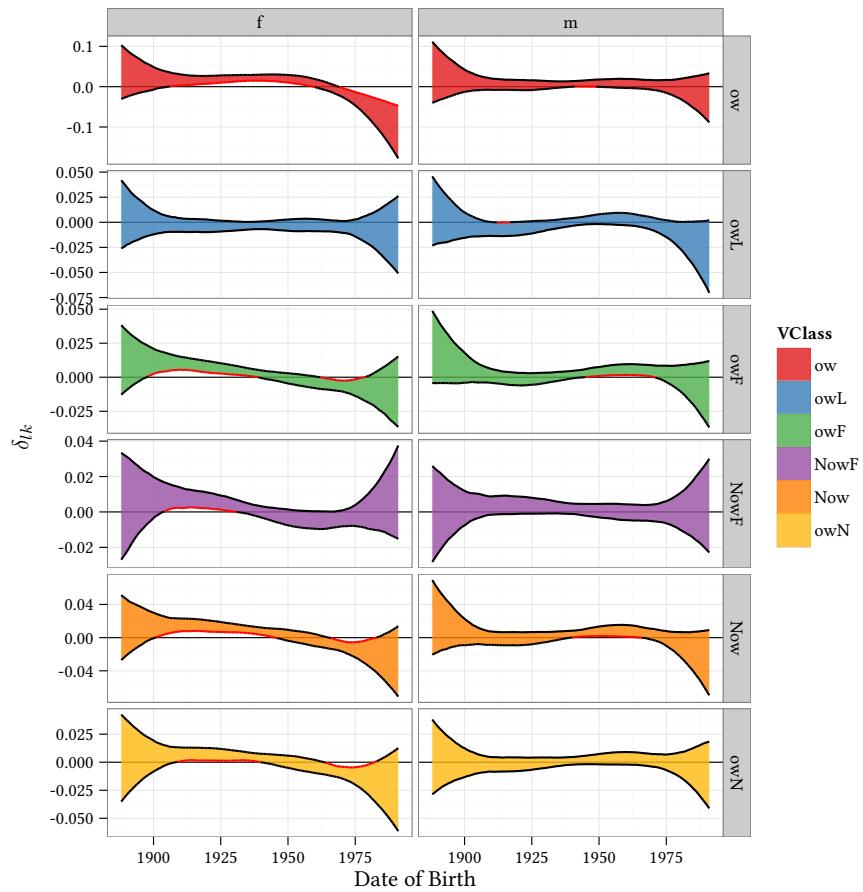


Figure 4.29: Year-to-year differences for variants of /ow/. Note: y-axis ranges differ across each horizontal set of facets.

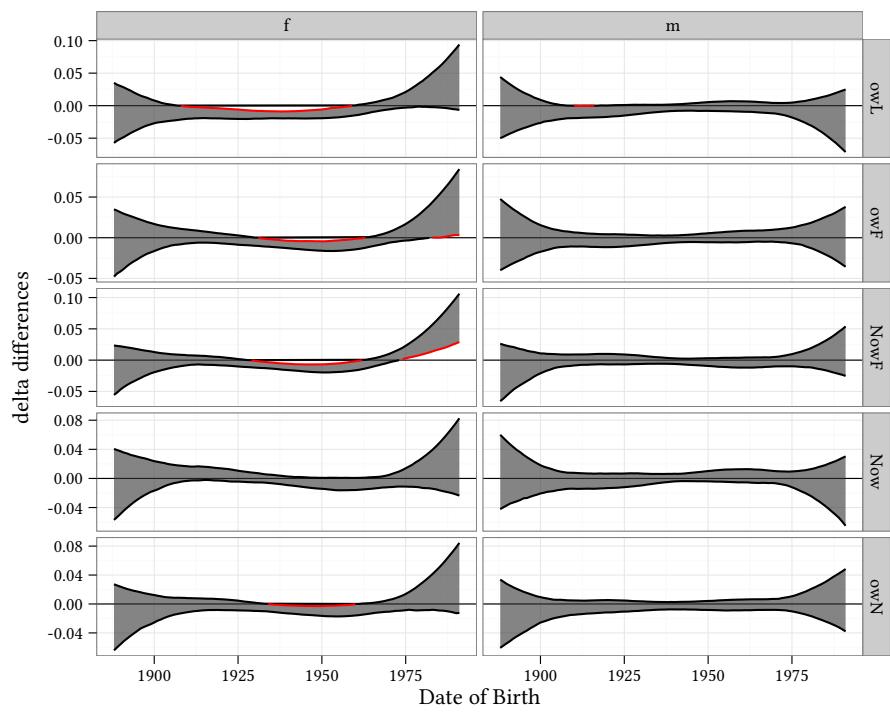


Figure 4.30: Rate of change differences from [ow]. Note: y-axis ranges differ across each horizontal set of facets.

exhibits a reliably different rate of change from [ow] is 1908. These dates are obviously overly precise, but support the interpretation where [ow] and [owL] have been categorically differentiated from the very beginning of /ow/ fronting.

Importantly, it is highly unlikely that the differentiation of [ow] and [owL] could be due to changing degrees of coarticulation between /ow/ and /l/. In this change, [ow] is undergoing a fronting change, and [owL] is being left behind. The estimated rate of change for [owL] (Figure 4.29) contains 0 across the entire century, meaning that if the difference between [ow] and [owL] were due to phonetic coarticulation, the strength of the coarticulation effect would have to be increasing exactly in proportion to the degree of frontness of [ow]. This is highly unlikely, and positing a categorical differentiation between [ow] and [owL] is the simpler explanation.

Surprisingly, even though [owF], [NowF] and [owN] share the same profiles as [ow] in both their overall trajectories and in their rates of change, there are some reliable differences between their rates of change. All four of these variants began fronting at approximately the same time around the turn of the century, but [ow] continued fronting until about 1960, while [owF], [NowF] and [owN] stopped fronting in the 1930's. Table 4.6 contains the model estimates for the dates when fronting began and ended, based on when the lower bound of the HPD for δ_{lk} excluded 0. [owF] and [NowF] also appear to be sluggish in participating in /ow/ retraction which began in the mid-1960's for [ow].

Variant	Began Fronting	Stopped Fronting
ow	1906	1959
owF	1898	1938
NowF	1904	1930
owN	1909	1939

Table 4.6: Dates that /ow/ variants began and stopped fronting, based on δ_{lk}

I think a reasonable analysis for [owF] and [NowF] is that we are observing a ceiling effect. Looking at the trajectories of change, [owF] and [NowF] in Figure 4.28 are the most fronted /ow/ variants, an effect that can be probably be simplified to simply being word final. If, for reasons which are unclear, there were a maximal degree to which /ow/ could be fronted phonetically, then it would make sense that [owF] and [NowF] would hit this limit first, and bottom out. In fact,

if we look at the trajectories of change for [ow], [owF] and [NowF], and examine the estimates for the maximum values of F2 these variants reached, we see that all three variants reached very similar peaks, even though the times at which they reached these peaks are spread out over 25 years. This suggests that there is, in fact, some ceiling value around $F2 \approx -0.5$ which [owF] and [NowF] hit first, slowing down their rates of change.

Variant	Maximum F2	Date of maximum
ow	-0.49	1964
owF	-0.56	1948
NowF	-0.61	1940

Table 4.7: Maximum /ow/ F2 values, and dates they were reached.

This ceiling analysis does not extend to [owN], however, which also stops fronting much earlier than [ow], but reaches a much lower peak value. The fact that [owN] could not possibly be slowing down due to a ceiling effect, and that its rate of change is reliably slower than [ow] starting in the mid 1940's, means that this may be a candidate example for the reanalysis of a phonetic effect as a phonological process.

Discussion

This analysis of /ow/ illustrates some interesting limitations of the rate of change diagnostic. First, in order to be able to diagnose anything at all, there must be a change occurring. In this case, the non-participation of men in /ow/ fronting meant that nothing can be said with much certainty about the phonetic and phonological status of contextual variants of /ow/ in their speech. Second, the analysis will be sensitive to ceiling and floor effects. Both [owF] and [NowF] phonetically favored fronter /ow/, and therefore reached the ceiling in the phonetic space (at approximately -0.5) first, which flattened out their trajectory of change, reducing their rate of change. I believe saying that the reduction in the rate of change of [owF] and [NowF] relative to [ow] is due to a ceiling effect is well founded, because [owF], [NowF] and [ow] all reached the same peak F2 value, but at different points in time. The fact that the rate of change for /owN/ slowed at a point which could not be considered a ceiling means that it should be held out as a potential case of a phonetic

bias becoming reanalyzed as a phonological process.

Finally, I will be analyzing the categorical exemption of /owL/ from fronting as a phonological distinction. For now, I will propose the process in (4.30) which I will further support in §4.5.

$$(4.30) \text{ } \text{ow} \rightarrow \text{o:/ } \underline{_} \text{l}$$

Under this analysis, only /ow/ which has a phonological glide target is affected by /ow/ fronting, while the long monophthong remains fully back.

4.4.3 /uw/

The third and final back-upgliding diphthong I'll be analyzing is /uw/. /uw/ has also undergone a fronting change which has also been reversed. As with /ow/, men have had very limited involvement in /uw/ fronting. Figure 4.31 displays the basic trajectories of /uw/ for men and women in F1×F2 space. As with /uw/, the primary acoustic dimension describing the change is F2, so the following models will focus on normalized F2.

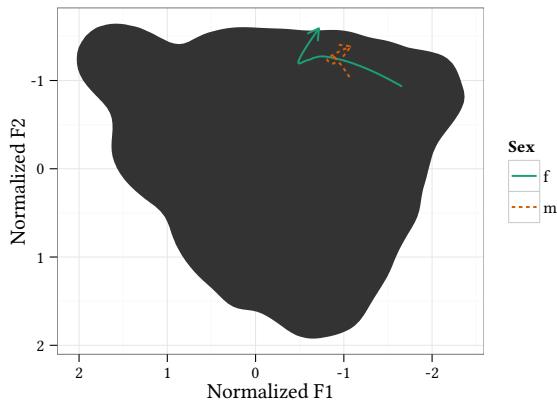


Figure 4.31: /uw/ Trajectory in F1×F2 Space

The two biggest conditioning factors on /uw/ fronting are following /l/, which favors backer /uw/, and preceding coronals, which favor fronter /uw/. The possible reasons for /l/ favoring backer /uw/ carry over from the discussion of /ow/. The effect of preceding coronals, however, is new. It is, in fact, a more general property of North American English that /uw/ tends to be fronter when preceded by a coronal, as discussed in the Atlas of North American English (Labov

et al., 2006, ch 12). The ANAE proposes that this phenomenon is related to the merger of /juw/ and /uw/ post-coronally in North America. For example, there are the differences between RP and “Standard American” presented in Table 4.8, with very broad transcriptions. The ANAE argument is that before /j/ was lost in this context in North America, it had the effect of fronting the nucleus of /uw/, which has persisted. Furthermore, /tu:/ sequences which never had a /j/ (e.g. *do* vs. *dew*) have merged to the fronted version. This is a classic phonologization argument, in the sense if Hyman (1976), and is briefly sketched out in Table 4.9.⁶

	RP	North American
<i>tube</i>	/tju:b/	/tu:b/
<i>tune</i>	/tju:n/	/tu:n/

Table 4.8: Examples of post coronal /j/ loss in North America

tju:b Initial state: little coarticulation	>	tji <u>b</u> Coarticulation of /j/ and /uw/	>	ti <u>b</u> Phonologization
---	---	--	---	--------------------------------

Table 4.9: The phonologization of post-coronal /uw/ fronting

Another possible account for the coronal effect on /uw/ fronting is that it is simply a case of coarticulation. Figure 4.32 is an illustration of the coarticulatory effect of coronals on back vowels from Ohala (1981). This illustration is of anticipatory coarticulation, where the coronal follows the vowel, but it could extend in principle to the case here where the coronal precedes the vowel.

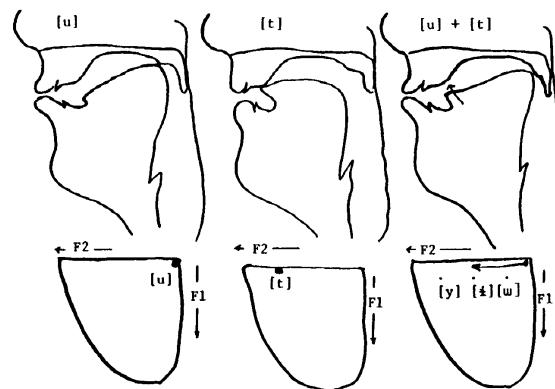


Figure 4.32: Illustration of the effect of coarticulation on /uw/ from Ohala (1981)

⁶A very similar account could be given the development of [tjub]>[tjub] for many British speakers.

The difference between the ANAE account of the coronal effect and Ohala's is, in fact, precisely the difference between phonological and phonetic conditioning that I would like to use the rate of change analysis to resolve.

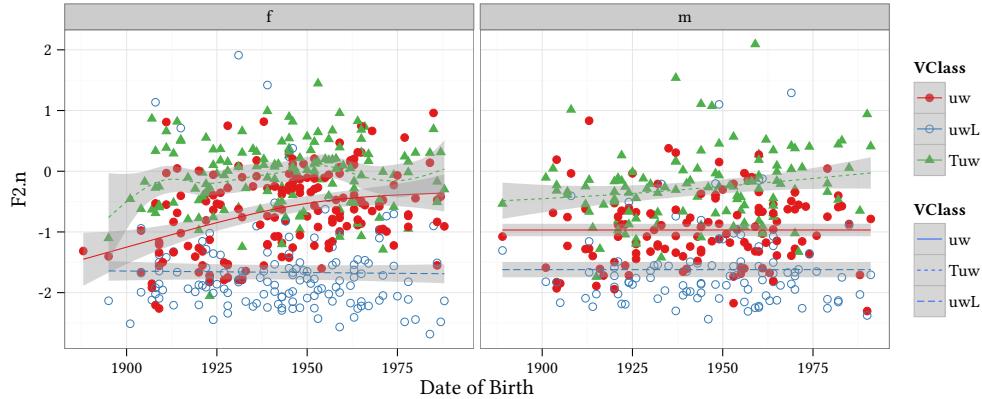


Figure 4.33: The effect of following /l/ and preceding coronal on /uw/.

Figure 4.33 displays the trajectories of the three basic variants to be investigated in this section. Again, it looks as if the [uwL] variant is categorically exempted from the change, but the pattern is more ambiguous for [Tuw]. Table 4.10 lists the coding criteria for /uw/ variants, and Table 4.11 displays the counts for each variant. I coded for whether /uw/ was followed by a nasal, /l/, or was word final, and for whether or not it was preceded by a coronal.

Variant	Criteria
TuwN	Is preceded by a coronal and followed by a nasal within the same word
TuwF	Is preceded by a coronal and is word final
TuwL	Is preceded by a coronal and followed by /l/ within the same word
Tuw	Remaining cases which are preceded by a coronal
uwN	Is followed by a nasal
uwF	Is word final
uwL	Is followed by /l/
uw	Remaining cases

Table 4.10: Coding criteria for /uw/

The most frequent variant, [TuwF], is composed mostly of the lexical items *do*, *to*, *two*, *too*, but also consists of a number of other lexical items, so it should be well modeled. The variants [uwN] and [TuwL] are too infrequent to model in this way, which is especially unfortunate for [TuwL],

Variant	N	Word Types
TuwF	9564	27
uw	2035	149
uwL	1474	31
Tuw	1461	180
uwF	883	24
TuwN	589	56
uwN	148	27
TuwL	78	16
total	16232	509

Table 4.11: Token counts of each /uw/ variant

since it could be crucial to see how a favoring and disfavoring context interact. Figure 4.34 plots the trajectories of the remaining variants which were fitted by the model. The reference variant in this model will be [uw], the variant which is not post-coronal, pre-nasal, pre-/l/, nor word final.

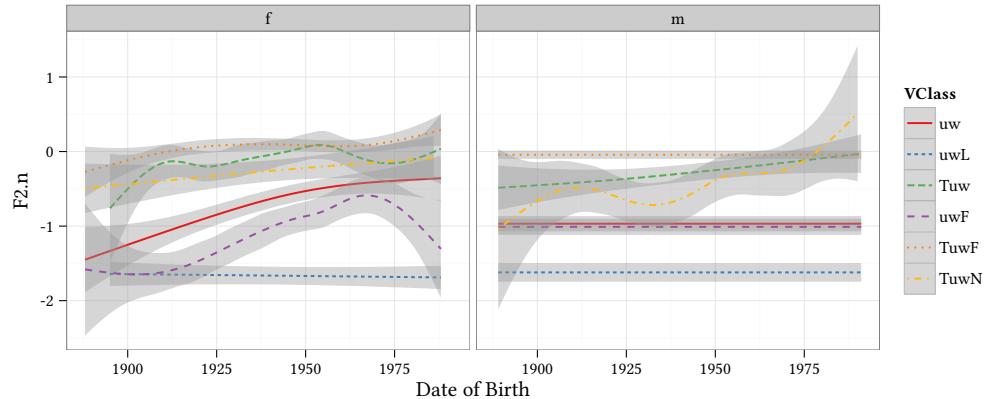


Figure 4.34: The /uw/ variants to be modeled.

Model Fit

Figure 4.35 displays the distribution of the \hat{R} convergence diagnostic for the estimated trajectory of /uw/ variants. [TuwF] and [uwF] have the largest \hat{R} values, but they are still acceptably close to 1 to include them in the analysis.

Figure 4.36 plots the estimated trajectories of /uw/ variants, following the same convention

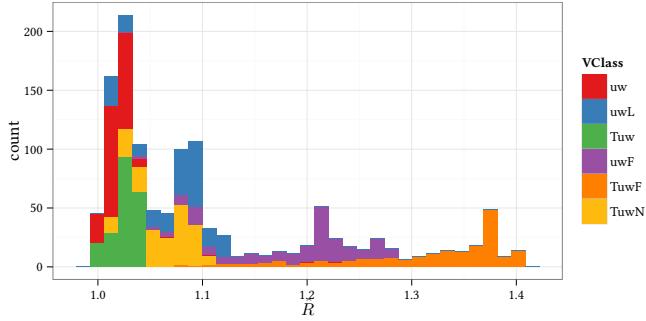


Figure 4.35: \hat{R} for the predicted trajectories of /uw/ variants.

used for /ow/, where the trajectory of the reference level, [uw], is superimposed over the trajectory of each other variant. The trajectories fitted by the model replicate the disfavoring and diverging effect of [uwL], as well as the favoring effect of a preceding coronal, which is stably in place regardless of the following segment, since [Tuw], [TuwF] and [TuwN] are all displaced upwards along F2. The sociolinguistic difference between men and women is also on display in Figure 4.36, where men appear to be only minimally participating in the change. Based on the results for /ow/, this will mean that the rate of change diagnostics for men will be only minimally informative.

Figure 4.37 plots the estimated rates of change for these /uw/ variants. [uw] has a positive rate of change starting at the turn of the century, as does [uwF] and most post-coronal variants. Before /l/, it looks like [uwL] has been completely flat over the course of the century, just like /ow/, and may have even undergone some retraction in the 1970s. A notable pattern for many of the post-coronal variants is a double dip, where they start out with a positive rate of change, level out, and then begin fronting again. There is a positive rate of change around the mid 1960s and 1970s for [TuwF] for women, and [Tuw] and [TuwN] for men, which is not present in either [uw] or [uwF].

Turning now the the crucial comparison of the rate of change of [uw] to the other variants, we see that [uwL] has had a reliably slower rate of change than [uw], unsurprisingly. A number of post-coronal variants also appear to have a period of time where their rate of change is slower

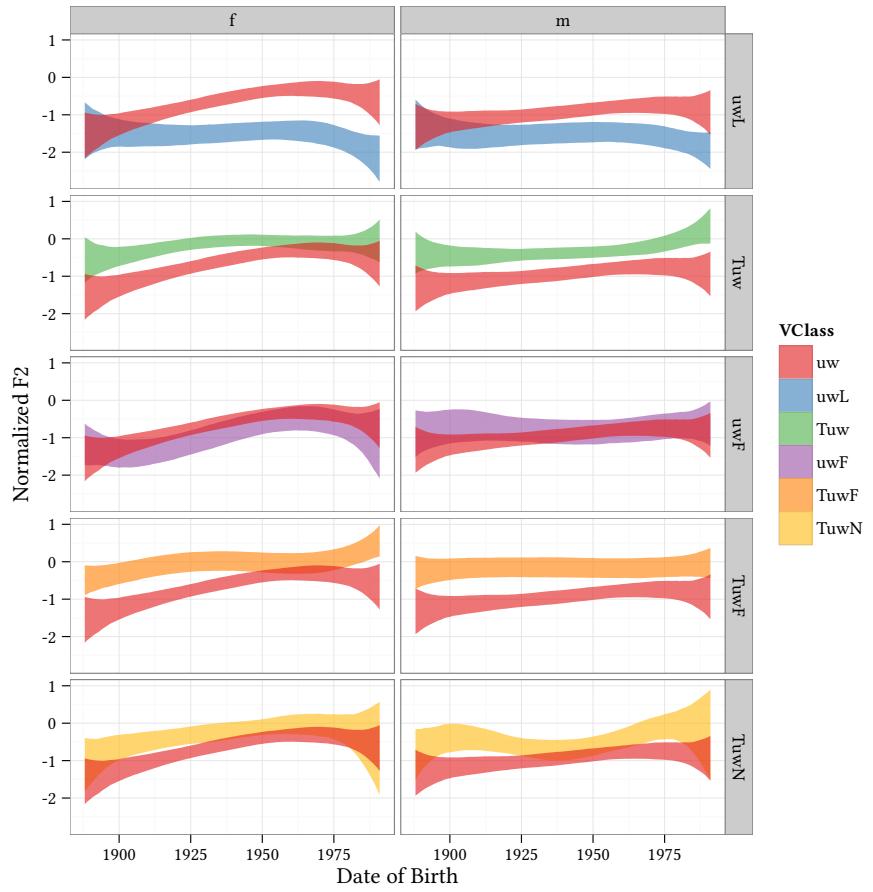


Figure 4.36: Predicted trajectories of change for /uw/ variants.

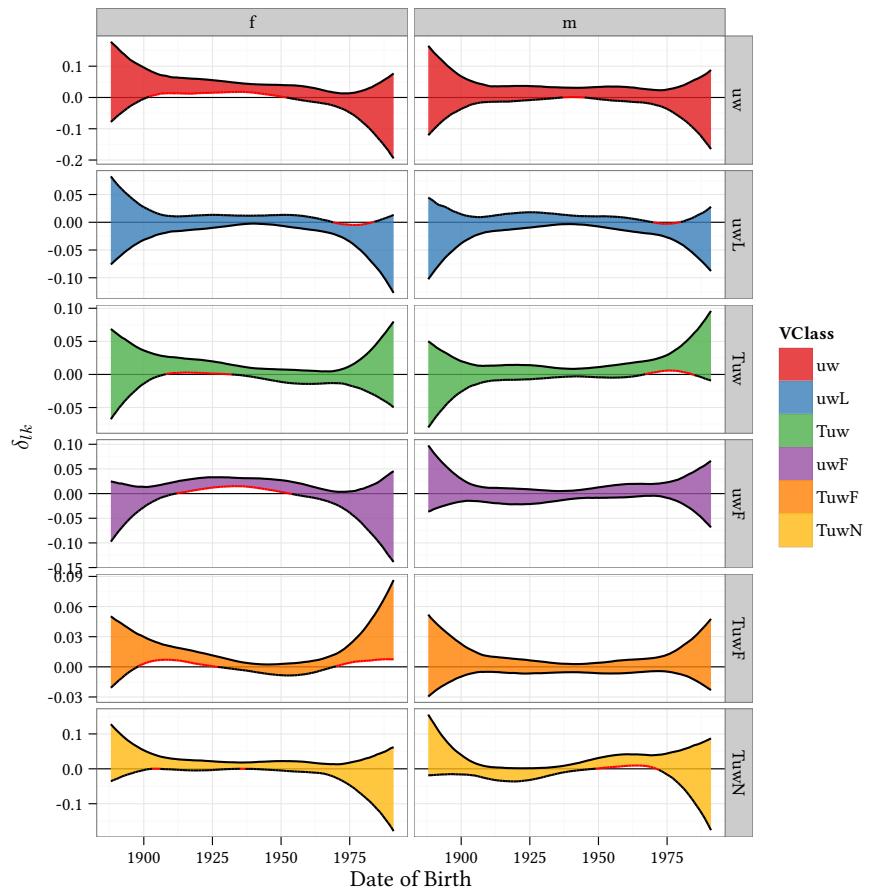


Figure 4.37: Year-to-year differences for variants of /uw/. Note: y-axis ranges differ across each horizontal set of facets.

than [uw]: [Tuw] and [TuwF] for women, and [TuwN] for men.

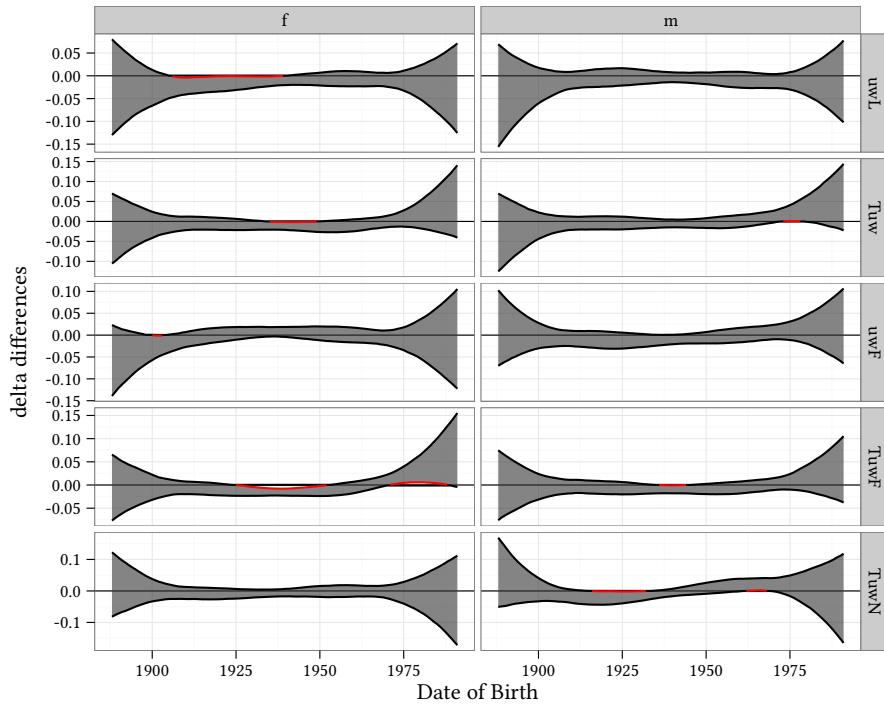


Figure 4.38: Rate of change differences from [uw]. Note: y-axis ranges differ across each horizontal set of facets.

Given that a preceding coronal favors /uw/ fronting, can we attribute the reliably slower rates of change for these post-coronal variants to a ceiling effect, like I proposed for word final /ow/? It is a possibility, but it seems less likely in this case. My argument for a ceiling effect on /ow/ rested on the fact that all of the /ow/ variants reached very similar maxima, but at different times. In the case of post-coronal /uw/, these variants had clearly not reached their maxima in the 1930s and 1940s, because in the 1960s and 1970s they began to front some more, as was seen in Figure 4.37. Perhaps this second phase of fronting for post-coronal /uw/ could be interpreted as being a phonological reanalysis of the coronal coarticulatory effect. If so, however, it occurred at the strangest time: when [uw] and [Tuw] were actually minimally different, as can be seen in Figure 4.36. It seems clear from the rate of change diagnostic that post-coronal /uw/ is phonologically distinguished from /uw/ in other contexts, but unfortunately, the timing of when this phonological differentiation entered the grammar relative to the phonetic fronting of /uw/ is ambiguous.

The effect of following /l/, on the other hand, is unambiguous. [uwL] has a reliably slower rate of change from [uw] almost as soon as [uw] begins fronting. Just like it was for /ow/, the pre-/l/ variant of /uw/ is being categorically exempted from ever undergoing the change.

Event	Year
[uw] is reliably fronting	1902
[uwL] is reliably slower than [uw]	1906

Table 4.12: Comparing the timing of [uw] fronting and the differentiation of [uwL]

4.5 Summary of /Vw/ results

Table 4.13 summarizes the results of the case studies just presented. The conditioning factors which were labeled in previous literature as having substantial effects are bolded. There is only one case which fits the profile of potential phonological reanalysis of a phonetic effect, [owN]. The rest of the unambiguous cases either exhibit parallelism throughout the change (with the exception of [owF], which is explicable by a ceiling effect), or were divergent from the very start of the change.

	/aw/	/ow/	/uw/
VwN	Always Parallel	Potential Reanalysis	–
NVw	–	Always Parallel	–
VwL	–	Divergent from start	Divergent from start
VwF	Always Parallel	Ceiling Effect	Always Parallel
TVw	–	–	Ambiguous

Table 4.13: Summary of /Vw/ Results

The possible source of the /l/ effect

The stand out effect here has been following /l/. For both /uw/ and /ow/, a following /l/ has categorically blocked fronting since the very beginning of the change. It is true that we would expect a following /l/ to disfavor the fronting of back vowels phonetically, especially in Philadelphia

where /l/ undergoes darkening and vocalization at a higher rate and in more environments than other North American dialects (Ash, 1982). But based on the argumentation in §4.1.1 and §4.2, the *categorical* nature of /l/ blocking means that /l/ must have had a categorical phonological effect on these vowels that was present in the grammar either before, or concurrent with the phonetic fronting of the other /ow/ allophones. In fact, I believe it is plausible to argue that this effect is due to glide simplification.

A rather salient feature of the Philadelphia dialect is that the glide of /aw/ is deleted when followed by an /l/. This is a rather old feature of the dialect, as it was mentioned by Tucker (1944), who said:

When *ou*, pronounced [æʊ], loses its second element, the result simply ‘flat a’: *hour* [æ:r], *owl* [æ:l], *Powell* [pæ:l], the latter two hardly to be distinguished from *Al* and *pal*.

In this same description of the Philadelphia dialect, Tucker also explicitly notes that Philadelphians make no distinction in the vowel quality of /ay/ before voiceless segments. The PNC data shows pre-voiceless /ay/ raising entering into the Philadelphia dialect with speakers born in the 1920s, meaning that we can reasonably place this process of glide deletion as being present in the dialect well before that. We could formulate the process of /aw/ glide deletion as follows.

$$(4.31) \quad \text{æw} \rightarrow \text{æ: } / __ l$$

This effectively captures the argument of Dinkin (2011a) that in Philadelphia /æl/ has merged with /awl/, leading to the raising and tensing of /æ/ before /l/. Dinkin (2011a) argues that rather than this being an extension of the Philadelphia split short-a pattern, this tensing and raising of /æ/ before /l/ is actually because its phonetic target is that of the nucleus of /aw/.

It would be reasonable to extend this process to cover both /ow/ and /uw/. To begin with, the phonetic realizations of the fully back /ow/ and /uw/ before /l/ is as long monophthongs: *bowl* [bo:u:l]⁷, *school* [sku:u:l]. Secondly, it is reasonable to categorize /aw/, /ow/ and /uw/ as belonging to a phonological natural class. The common patterning of these vowels cross-dialectally was the

⁷I'm using [u:l] to represent the extreme reduction of /l/ to a velar approximant. /l/ is not always vocalized in these contexts, and the coda can be occasionally accompanied by light velar frication.

motivating force behind the Labovian transcription conventions for these vowels, which I have largely adopted, and in Philadelphia, we observe all three undergoing a simultaneous fronting and reversal in parallel, a phenomenon which will be covered in depth in later chapters. The extension of glide deletion to /ow/ and /uw/ could be formulated as follows.

$$(4.32) \quad Vw \rightarrow V:/__l$$

This could be reformulated in moraic terms. We start out with an underlying form like (4.33). The /Vw/ glide is then delinked, triggered by the following /l/ (4.34). This may be an OCP effect of some sort, especially if /l/ is really a velar approximant in this position. Crucially, the mora originally associated with the glide becomes associated with the nucleus, creating a long monophthong (4.35). This would have the effect of exempting /ow/ and /uw/ from fronting in this context, because fronting only affects the nucleus of these vowels, i.e. the first mora.

$$(4.33) \quad \begin{array}{c} \mu & \mu & \mu \\ | & | & | \\ V & w & l \end{array}$$

$$(4.34) \quad \begin{array}{c} \mu & \mu & \mu \\ \backslash & / & | \\ V & w & l \end{array}$$

$$(4.35) \quad \begin{array}{c} \mu & \mu & \mu \\ \diagdown & | & | \\ V & & l \end{array}$$

4.5.1 Connection to Broader Theory

I am proposing that the glide deletion process discussed above was categorical, and must have been present in speakers grammars at the very start of the phonetic change that fronted /ow/ and /uw/ in all other contexts in order to categorically block fronting. Since neither [uwL] nor [owL] ever underwent fronting, and since the phonetic difference between [uw]~[uwL] and [ow]~[owL] was still very small at the time that categorical blocking was in place, this phonological process was not the reanalysis of phonetic coarticulation. This supports my general argument that phonetic change operates over the representations produced by a distinct phonology, and

that theories of sound change based solely in phonetics are insufficient to capture the facts of all, or most sound change.

4.6 Conclusion

In this chapter, I have laid out my definition of a phonetic effect, or phonetic coarticulation, in contrast to phonological differentiation, and examined how these different phenomena ought to interact with sound change. Importantly, my definition of an effect being phonological or phonetic is based on which domain of the sound system the effect originates, *not* on its size. It is possible under the model of the phonology-phonetics interface I have adopted for effects originating in the phonetics to be large, and produce discrete non-overlapping distributions, and for effects originating in the phonology to be small, and produce partially overlapping distributions. Before even examining empirical case studies, the phonology-phonetics interface model first described in Chapter 2, and fleshed out in more detail in this chapter, when combined with diachronic change, produces what I'll call the “Unity Principle,” and is very similar to what Kroch (1989) called the “Constant Rate Effect.”

(4.36) THE UNITY PRINCIPLE

If two contextual variants have the same surface phonological representation, then they must shift in parallel diachronically. Contrapositively, if two contextual variants have divergent diachronic trajectories, they must not have the same surface phonological representation.

The Unity Principle is not itself a falsifiable hypothesis, but rather a logical consequence of the phonology-phonetics model I have adopted. It serves as a tool for investigating the interaction of phonology and phonetics over the course of phonetic change.

Applying the Unity Principle to the fronting of /aw/, /ow/ and /uw/ was successful in terms of discovering new details about these particular changes and the broader generalizations they imply, as well as demonstrating the utility of the Unity Principle for phonological investigation. Some conditioning effects, like that of following nasals on /aw/, appear to be strictly phonetic,

despite for their large effect size, because they move in parallel over the entire course of the change. Many apparent exceptions to parallelism are reasonably understood in terms of ceiling effects. The biggest exception to parallelism was for /ow/ and /uw/ when followed by /l/. These variants appeared to be categorically exempt from fronting.

Faced with the divergent trajectories of [ow ~ owl] and [uw ~ uwL], it follows from the unity principle that these variants must have different surface phonological representations, meaning we must posit either a phonological process to differentiate them, or a different underlying form. The fact that it is *necessary* to posit a phonological analysis in the face of the diachronic data combined with the Unity Principle speaks to both its utility, and the importance of diachronic data for phonological investigation. Any given snapshot of [ow ~ owl] [uw ~ uwL] based on a demographically narrow set of speakers would be ambiguous, and open for reasoned argument for either a purely phonetic or purely phonological explanation. It is the diachronic dimension in combination with a moderately articulated model of the phonology-phonetics interface which disambiguates the two sources of explanation, and opens the door for more detailed inquiry.

The more surprising result with broader implications for language change in general is the relative timing of the phonetic change which began fronting [ow] and [uw], and when the phonological process differentiating [owl] and [uwL] must have been in the phonological grammar. Rather than the slow and gradual reanalysis of coarticulated [owl] and [uwL] as being their own phonological allophones over the course of the change, they behaved as categorically distinct allophones from the very beginning of the change. This effect of early phonological differentiation in phonological change will reappear in Chapter 5, and is unexpected under most accounts of conditioned sound change where the introduction of a phonological process follows a period of accumulated phonetic errors (Ohala, 1981; Blevins, 2004). Even Janda and Joseph (2003) who propose a “Big Bang” model of phonologization of sound change include a “brief” period of purely phonetic conditioning. The results presented in this chapter and at the beginning of Chapter 5 suggest that if there is a brief period of pure phonetic conditioning, it is too brief to be detectable by statistical methods. In fact, the available data is equally consistent with phonologization occurring *simultaneously* with the onset of the phonetic change.

Chapter 5

Phonologically conditioned divergence and convergence

In this chapter, I will further support the argument that neogrammarian phonetic change targets *phonological* categories through more detailed analyses of /ay/, and /ey/ raising, /aw/, /ow/ and /uw/ fronting, and /ae:/ and /ɔ:/ lowering. First, I will show differential participation in phonetic change for variants within vowel categories which is best explained in terms of phonological allophony, rather than phonetic predisposition. Second, I will show convergent participation in phonetic change across multiple categories, which can be best explained as phonetic change targeting a phonological feature which defines a phonological natural class.

5.1 Phonologically divergent behavior within categories

5.1.1 /ay/ Raising and Opacity.

I didn't include pre-voiceless /ay/ raising in the rate of change analysis because even if I found that pre-voiceless /ay/ had a divergent rate of change from other /ay/ (which it undoubtedly would), it would still be ambiguous between error accumulation, or increasing coarticulation, and phonological differentiation. Unlike /ow/ and /uw/, where most contexts were undergoing the change with some contextually restricted variants exempted, for /ay/ raising, the change takes place only for contextually restricted variants. If this contextual effect were due to phonological selection, it would be indistinguishable from a gradually increasing coarticulation.

Fortunately, though, voicing neutralization of /t, d/ by flapping occurs in Philadelphia, and

provides the ideal environment for distinguishing between phonetic and phonological conditioning of /ay/ raising. In contemporary Philadelphian English, /ay/ raising applies opaquely with respect to flapping, producing the same dilemma identified by Joos (1942) in Canada.

Table 5.1 provides an ordered rules analysis⁸ of /ay/ raising in contemporary Philadelphian English.

	<i>writer</i>	<i>rider</i>
Input	ɹaɪtə	ɹaɪtə
RAISING	ɹaɪtə	-
FLAPPING	ɹaɪɾə	ɹaiɾə
Output	ɹaɪɾə	ɹaiɾə

Table 5.1: Opaque interaction between /ay/ raising and flapping.

This opacity represents an important end point for the process of /ay/ raising. If we were to assume that /ay/ raising began as a coarticulatory process, then there *must* be a point in its history when it became reanalyzed as a phonological process conditioned on the underlying voicing of the following segment. And, as the rate of change analysis in the previous chapter indicated, this point of reanalysis must happen within the time period covered by the PNC, since the onset of the change in the first place appears to be contained within the PNC.

Foundational Facts

There are some more foundational facts about /ay/ raising that should be established before delving into its interaction with flapping. Specifically, I should establish how pre-voiceless /ay/ raising is conditioned in Philadelphia in contrast with previous descriptions in other dialects. Dailey-O'Cain (1997), for instance, describes /ay/ raising as applying before /r/ as well as pre-voiceless, and Idsardi (2007) reports his intuition that it can apply across word boundaries.

Figure 5.1 plots /ay/ height over date of birth as conditioned by manner and voicing of the following segment. It is clear from this figure that it is only voicing which conditions raising. While there is some considerable contextual variation in height in the following voiced contexts, it appears to be less extreme within the following voiceless context.

⁸Intended merely for expository purposes.

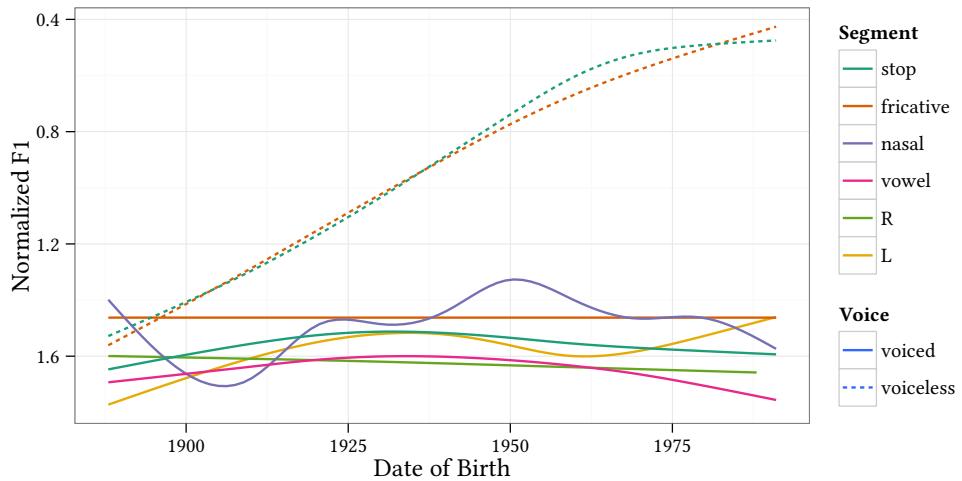


Figure 5.1: The effect of following voice and manner on /ay/ height.

Most importantly, there is no tendency for /ay/ to raise when followed by /r/, distinguishing the phonological system of Philadelphia from the Northern dialects which do raise /ay/ before /r/, as reported by Dailey-O'Cain (1997). This difference is suggestive of the fact that pre-voiceless /ay/ raising is an endogenous change to Philadelphia, not a dialectal borrowing. The process of dialectal diffusion can lead to structural simplifications, but in this case there is no independent historical reason to assume that /ay/ raising diffused from Northern dialects to Philadelphia, as there was for the diffusion of the New York City short-a system to Cincinnati and New Orleans (Labov, 2007).

Next, looking at word final /ay/, there does not appear to be any conditioning of /ay/ raising by the onset of the following word. The trajectories of word final /ay/ when followed by voiced and voiceless onsets are virtually identical. Even if we take a targeted subset of word final /ay/ followed by *to* and *the*⁹ there is still no evidence of /ay/ raising being conditioned across word boundaries. In contrast to the Canadian dialect from which the intuitions reported by Idsardi (2007) were drawn from, there is no evidence that pre-voiceless /ay/ raising was ever a phrase level phenomenon in Philadelphia. This in itself is already evidence for phonological conditioning of

⁹The motivation here is to focus on prosodically weak words which are more likely to “lean” on the preceding /ay/.

/ay/ raising, since raising is sensitive to word boundaries, and word boundaries are a phonological property.

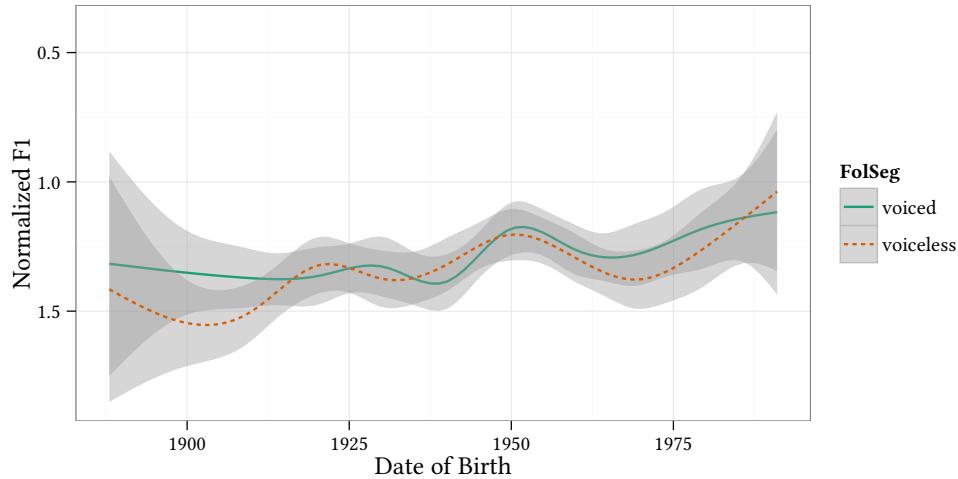


Figure 5.2: Effect of following word onset on word-final /ay/.

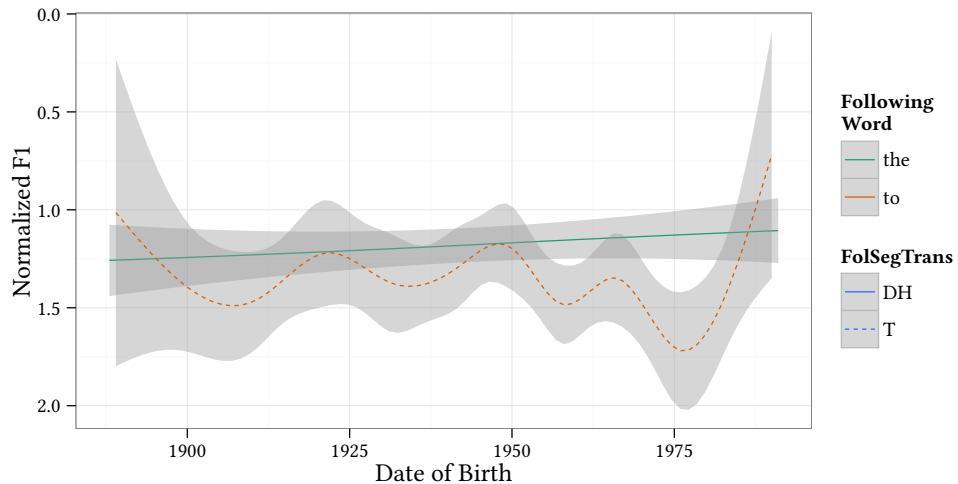


Figure 5.3: Effect of following *to* and *the* on word-final /ay/.

Interaction with flapping

It is necessary to construct two careful subsets of the /ay/ data in order to investigate the history of the interaction between /ay/ raising and /t, d/ flapping. One is a subset where /ay/ appears before

/t, d/ which are almost certainly not flapped. I will be referring to this subset as “surface” /t, d/, meaning that the underlying voicing contrast is realized on the surface. The second is a subset which appears before almost exclusively flapped /t, d/. I’ll be referring to this subset as “flapped” /t, d/. I defined these subsets as follows.

- (5.1) SURFACE: /ay/ followed by /t, d/ which are then followed by a pause, labeled sp in the forced alignment transcriptions.
- (5.2) FLAPPED: /ay/ followed by /t, d/ which are then followed by an unstressed vowel within the same word.

I decided to restrict the surface subset to require that the following /t, d/ be followed by a pause to avoid any ambiguities introduced by phrase level flapping, or any other phrase level processes. Occasionally, the aligner will mistake an exceptionally long /t, d/ closure as a pause, but this kind of error is still acceptable for my purposes here, since a /t, d/ closure long enough to be labeled a pause will certainly not be resulting from flapped variants of /t, d/. For the flapping subset, the onset of an unstressed syllable meets the structural description for flapping to occur. I haven’t inspected these tokens to make sure they are actually flaps, but as I’ll show in the duration domain, the difference between /t/ and /d/ appears to be mostly neutralized. Of course, word final /t, d/ can also flap when followed by a word with a vowel onset, but I decided that including /t, d/ in this context would include a mix of flapped and surface /t, d/, and an auditory inspection of all these tokens would be necessary.

I also defined the following exclusions from the flapping subset as defined above.

- (5.3) Potential glottalizing contexts. e.g. /t/ followed by syllabic /n/, as in *frighten* [frʌɪ?n̩]
- (5.4) Exceptional raising words, as identified in Fruehwald (2008). e.g. *spider, Snyder*

A justified criticism to the second of these exclusions is that I’ll have excluded exactly those cases which run counter to my hypothesis. However, examining these cases separately, it appears as if the exceptional raising in these words is a later development. In the entire PNC, there are 60 tokens of *Snyder*, 12 of *spider* and 1 of *cider*. Figure 5.4 plots the mean F1 for these words for each

speaker, and contrasts them with the height of /ay/ before flapped /t, d/ for those same speakers who contributed exceptional raising words. On average, these exceptional raising words did not appear raised until approximately 1940, about 20 to 30 years after pre-voiceless raising began in the dialect. Separating out these items as being reflective of a later development in the dialect is therefore principled. Table 5.2 displays the counts of observations in each context after taking these subsets and exclusions.

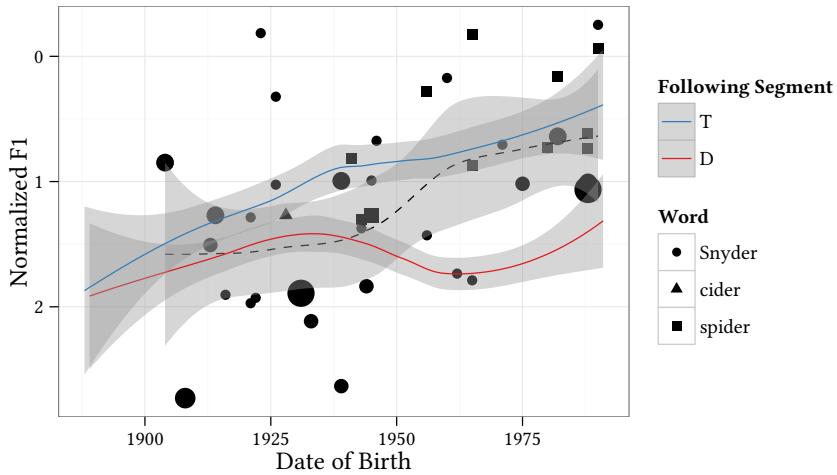


Figure 5.4: Exceptional raising words compared to /ay/ before flapped /t, d/ in all other words.

	/t/	/d/
Surface	2155	647
Flap	240	328

Table 5.2: Number of /ay/ observations in each context

Since my goal is to be able to disambiguate between phonetic and phonological conditioning of /ay/ raising, I'll first assess the phonetic pressures for /ay/ raising. Pre-voiceless vowel shortening is the most commonly appealed to phonetic precursor for pre-voiceless /ay/ raising, starting with Joos (1942) and Chambers (1973). The argument is very similar to that presented for /u/ fronting in Ohala (1981): the diphthong /ay/ involves a long gesture across articulatory space, and when the vowel is shortened before voiceless segments, this gesture must be made in a compressed amount of time. In compensation, it is argued that speakers may raise the nucleus of /ay/ to reduce the

gesture length. Moreton and Thomas (2007) make some very cogent arguments against the pre-voiceless shortening account. They point out that in dialects which monophthongize /ay/, the monophthongization is least advanced before voiceless segments, meaning that in these dialects /ay/ has the the longest articulatory gesture in the context where it is supposed to be the most difficult according to the shortening account of raising. Their alternative hypothesis is that the glide is peripheralized in pre-voiceless contexts, capturing both the coarticulatory pressure to raise the nucleus towards the glide, and the resistance to monophthongization. It's not clear that either hypothesis can account for raising /ay/ before /r/ in the Inland North, since /ay/ is both relatively long before this sonorant, and it is one of the most favorable contexts for /ay/ monophthongization. However, since /ay/ does not raise before /r/ in Philadelphia, this may be beyond the scope of relevance for the data from Philadelphia.

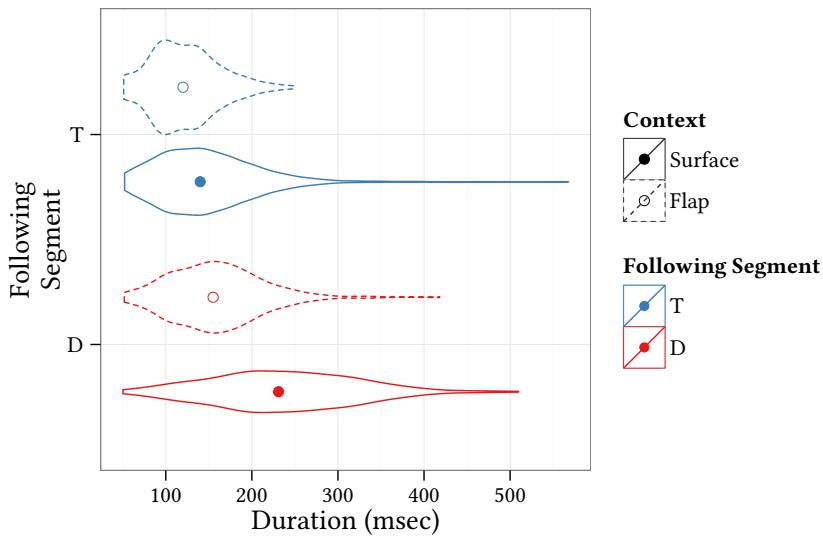


Figure 5.5: Violin plot representing the distribution of durations of /ay/ before surface and flapped /t/ and /d/.

I'll address these two hypotheses for the phonetic conditioning of /ay/ raising in turn, beginning with the pre-voiceless shortening hypothesis. Figure 5.5 is a violin plot representing the duration of /ay/ in the relevant contexts. The violin shape represents the density estimate, and the point within each violin represents the overall median. Based on the distribution in the figure, we can see that vowel duration is incompletely neutralized towards the shorter duration range before

flaps. Table 5.3 displays /ay/ contexts from shortest to longest.¹⁰ Since the PNC data was collected for the purpose of vowel formant measurements, vowels with durations shorter than 50ms were excluded. Additionally, the force aligner only has a duration resolution of 10ms. Because the data was censored at 50ms, and because the instrument of measurement does not have fine grained enough resolution, it would be inappropriate to attempt too much statistical inference of vowel duration as a response variable. However, we can see that over all, /ay/ before flaps and /ay/ before surface /t/ form one set of distributions, and /ay/ before /d/ forms a separate distribution.

Segment	Context	Median Duration (msec)	Difference from next shortest
/t/	flapping	111	–
/t/	surface	144	34
/d/	flapping	156	11
/d/	surface	237	82

Table 5.3: Median /ay/ durations by context.

If we were to assume that /ay/ raising is phonetically conditioned, and that the relevant phonetic conditioning is duration, then we should predict that flapped /t/, surface /t/ and flapped /d/ should all participate in raising. Even though the duration difference in /ay/ before flapped /t/ and /d/ is not completely neutralized (in the PNC data by a larger margin than recently reported by Braver (2011)), the duration of /ay/ before flapped /d/ is approximately the same as before surface /t/. It is uncontroversially established that /ay/ before /t/ undergoes the raising change, and if what really matters is the phonetic properties of the pre-/t/ context, we should expect other contexts with similar properties, like before flapped /d/, to also undergo the change. The fact that /ay/ raising ultimately ends up in an opaque relationship with flapping would necessitate a later reanalysis of raising as being phonologically conditioned, which would result in a trajectory of change which would look something like Figure 5.6.

¹⁰These medians were calculated in a three step process.

- (i) Calculate median /ay/ duration for each word within each speaker.
- (ii) From the medians in (i), calculate median /ay/ duration for each speaker in each context.
- (iii) From the medians in (ii), calculate median /ay/ duration for each context.

This three step process partially mitigates the imbalanced distribution of observations across speakers and lexical items.

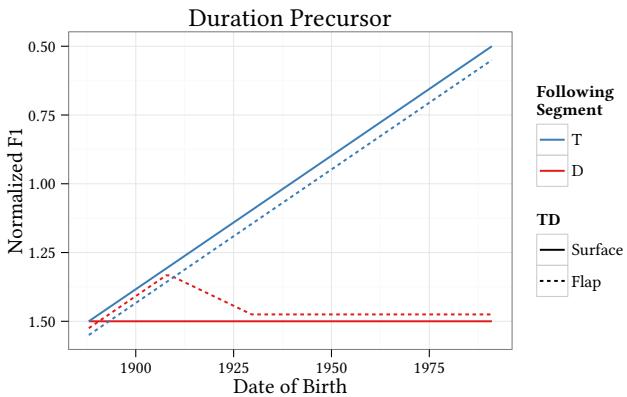


Figure 5.6: Schematic illustration of the reanalysis of /ay/ raising from being phonetically conditioned to being phonologically conditioned.

As for the glide peripheralization hypothesis, unfortunately glide measurements are not currently part of the PNC, and the results from other studies are somewhat inconclusive. Figure 5.7 plots data derived from Figures 4.1 and 4.2 of Rosenfelder (2005), a study of /ay/ and /aw/ raising in Victoria, British Columbia. The solid lines represent the average trajectories for /ay/ before all voiced and voiceless obstruents. Consistent with the glide peripheralization hypothesis, the glide targets (represented by the arrow heads) are much more peripheralized before voiceless obstruents than before voiced obstruents to a degree that is not in proportion to the difference in height of the nucleus. The dashed lines represent trajectories of /ay/ before flapped /t/ (trajectories for /ay/ before flapped /d/ were not reported separately in Rosenfelder (2005)). The glide target for /ay/ before flapped /t/ is much more similar to the glide target of voiced obstruents than voiceless obstruents. The difference in glide targets between pre-flap /ay/ and pre-voiced /ay/ appears to be more or less in proportion to the difference in nucleus height. Kwong and Stevens (1999) did a small scale acoustic study of pre-flap /ay/, and found that there was a statistically reliable difference in glide peripherality between pre-/t/-flap and pre-/d/-flap /ay/, but they did not include nucleus measurements, so it is impossible to tell if this difference is proportional to nucleus height differences. Moreover, Kwong and Stevens (1999) do not provide any /ay/ glide measurements from surface /t, d/ contexts, so it is also not possible to tell if the glide targets

before flaps pattern similarly across phonological categories, or if they are partially neutralized towards the pre-voiced glide targets as they appear to be in Rosenfelder (2005). Since Rosenfelder (2005) provides the relevant contrasts, the best interim assumption would be that before both /t/ and /d/ flaps, /ay/ glide targets are more similar to the targets of voiced obstruents.

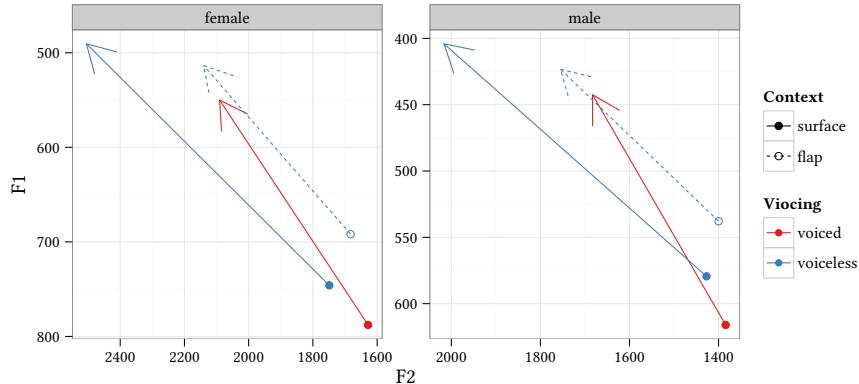


Figure 5.7: Nucleus to glide trajectories in Victoria, B.C. Derived from Rosenfelder (2005) figures 4.1 and 4.2

The glide-peripheralization hypothesis makes a very different prediction from the duration hypothesis. If voiced obstruents and both /t/ and /d/ flaps have similar glide targets, and other voiceless obstruents have peripheralized glide targets, it would predict that only surface /t/ should undergo raising. Again, the fact that /ay/ raising ultimately results in an opaque interaction with flapping means that if /ay/ raising were conditioned by glide peripheralization at first, it would eventually have to become reanalyzed as being phonologically conditioned, as illustrated in Figure 5.8.

Table 5.4 summarizes the difference between these two hypotheses in terms of which segments are predicted to undergo raising on the basis of phonetic conditioning. Under the duration precursor hypothesis, /ay/ preceding the set {surface /t/, flapped /t/, flapped /d/} should undergo raising, while /ay/ preceding {surface /d/} will not. Under the glide peripheralization precursor hypothesis, only /ay/ preceding the set {surface /t/} will undergo the change while /ay/ preceding {flapped /t/, flapped /d/, surface /d/} will not.

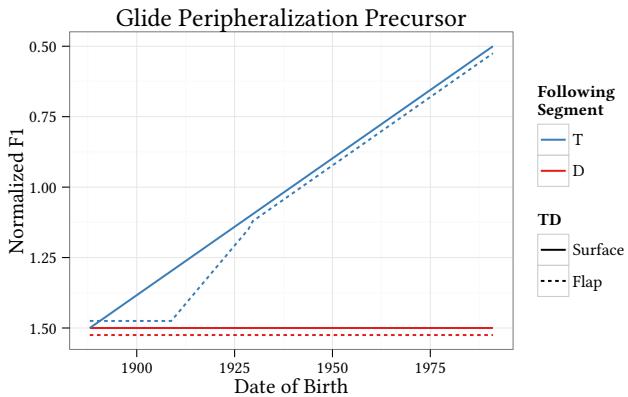


Figure 5.8: Schematic illustration of the reanalysis of /ay/ raising from being phonetically conditioned to being phonologically conditioned.

Duration Hypothesis	Raises			Low
	surface /t/	flapped /t/	flapped /d/	
Glide Peripheralization Hypothesis	Raises			Low

Table 5.4: Defining the undergoers and non-undergoers phonetically according to two different precursor hypotheses.

Having established the predictions of these two phonetic conditioning hypotheses, we can examine the actual data. Figure 5.9 plots cubic regression splines over speaker means for /ay/ in the contexts under discussion. The trajectories in Figure 5.9 don't match the predictions of either the duration precursor hypothesis or the glide peripheralization hypothesis. Rather, across the 20th century, the height of /ay/ appears to pattern according the underlying *phonological* voicing of the following segment, with {surface /t/, flapped /t/} patterning together, and {surface /d/, flapped /d/} patterning together.

Of course, the importance of phonological voicing to /ay/ raising is merely a qualitative impression of Figure 5.9 which will be supported by statistical inference in the next section. The implications are that /ay/ raising has always been a phonologically conditioned process, and that this phonological process has always interacted with flapping opaquely. This would require the addition of a phonological process to the grammar at the *onset* of the phonetic change, which at

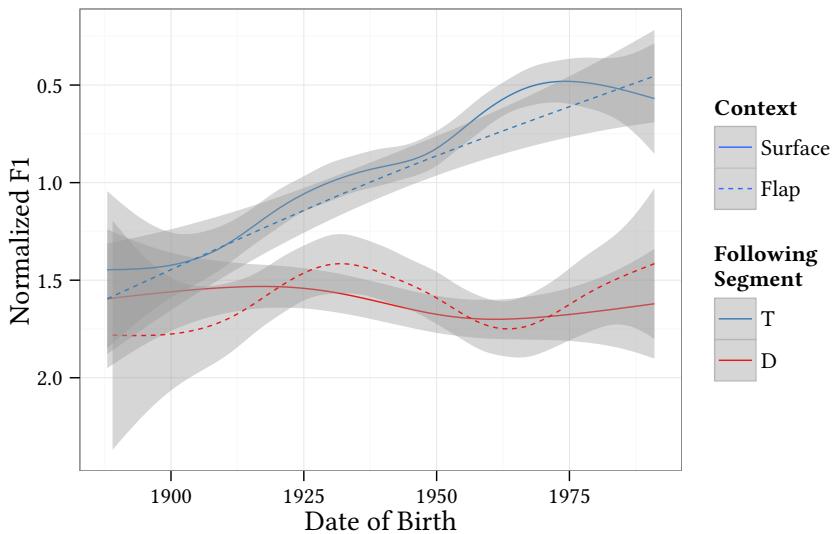


Figure 5.9: /ay/ height by date of birth and context. Cubic regression splines are fit over speaker means.

first corresponds to a small phonetic difference.

Modeling

To investigate the interaction of /ay/ raising and flapping more precisely, I constructed a hierarchical Bayesian model similar to that used in Chapter 4. The goal of this model is to determine whether there was ever a period of /ay/ raising where /ay/ allophones did not pattern according to the underlying voicing of following segment. There are at least two different ways we can try to answer this question.

- (5.5) Is there any point in time where the height of /ay/ is different between flapped and unflapped versions of the same segment?
(e.g. /ay/ could be a little bit lower before flapped /t/ than before unflapped /t/).
- (5.6) Is there any point in time when the difference in height of /ay/ between a following /t/ and /d/ is smaller when the /t/ and /d/ are flaps than when they are not flapped.

(e.g. There could be no height difference across the flaps, while there is a reliable difference between non-flapped /t/ and /d/).

As with the rate of change models from Chapter 4, I will be modeling functions over dates of birth using b-splines. See Chapter 4 §4.3.1 to review of how b-splines work. For this model, there are four variables that I want to model as a function of date of birth:

(5.7) The height along F1 of /ay/ before surface /d/.

(5.8) The difference in height for /ay/ between surface /d/ and surface /t/.

(5.9) The difference in height for /ay/ between surface /d/ and flapped /d/.

(5.10) The difference in height for /ay/ between surface /t/ and flapped /t/.

Each of these functions will be represented by γ_{le} , where l is the date of birth, and e is an index for the specific function. Each γ_{le} will be modeled using a b-spline.

$$\gamma_{le} = b.\text{spline}(l) \quad (5.11)$$

Now, with the exception of $\gamma_{l,1}$, none of these γ_{le} functions model the specific height of /ay/ in a given context. Instead, the community level estimate for the height of /ay/ in a specific context for a specific date of birth will be given as f_{lpk} , where l is the date of birth, p is the phoneme, and k is for the context.

$$f_{lpk} \quad (5.12)$$

$$p = \{D, T\} \quad (5.13)$$

$$k = \{\text{surf}, \text{flap}\} \quad (5.14)$$

The relationship for each height function f_{lpk} to the γ_{le} functions is given as follows.

$$f_{l,D,\text{surf}} = \gamma_{l,1} \quad (5.15)$$

$$f_{l,T,\text{surf}} = \gamma_{l,1} + \gamma_{l,2} \quad (5.16)$$

$$f_{l,D,flap} = \gamma_{l,1} + \gamma_{l,3} \quad (5.17)$$

$$f_{l,T,flap} = \gamma_{l,1} + \gamma_{l,2} + \gamma_{l,4} \quad (5.18)$$

As a consequence, each γ_{le} function can be interpreted as follows.

(5.19) $\gamma_{l,1}$

- The height of /ay/ before surface /d/ over time.

(5.20) $\gamma_{l,2}$

- The height difference from /ay/ before surface /d/ to /ay/ before surface /t/ over time.

(5.21) $\gamma_{l,3}$

- The height difference from /ay/ before flapped /d/ to /ay/ before surface /d/.

(5.22) $\gamma_{l,4}$

- The height difference from /ay/ before flapped /t/ to /ay/ before surface /t/.

For the sake of clear understanding, I won't be labeling axes in the figures which follow with the specific γ_{le} label, but rather these descriptions. Of particular interest will be whether $\gamma_{l,3}$ (the difference in /ay/ between surface and flapped /d/) and $\gamma_{l,4}$ (the difference in /ay/ between surface and flapped /t/) ever exclude 0 at any point in time. If they do, then it would mean that for that particular point in time, the effect of the following segment on the preceding /ay/ was not equivalent between surface and flapped forms.

Also of interest will be whether the difference in height between /t/ and /d/ is always the same whether they are flapped or not. We can represent this estimate as δ_l .

$$\delta_l = (f_{l,D,surf} - f_{l,T,surf}) - (f_{l,D,flap} - f_{l,T,flap}) \quad (5.23)$$

If δ_l ever excludes 0, that would mean that for that period of time, the difference in /ay/ height before /t/ and /d/ is different depending on whether they are flapped.

Now, one possible criticism of the model as I've laid it out so far is that I've built in the assumption that I'm trying to test. The height of /ay/ before surface /t/ is derived from the height

of /ay/ from surface /d/ plus $\gamma_{l,T,surf}$. The height of /ay/ before flaps are then derived from the height of /ay/ before the surface realizations. However, the functions for γ_{lpe} are only very weakly biased towards 0. If there was not sufficient data to estimate, say, the function for $\gamma_{l,T,flap}$ (the difference in height for /ay/ between surface /t/ and flapped /t/), its posterior distribution would be only marginally different from its prior, which in this model was $\mathcal{N}(0, 1000)$, meaning that values ranging from -1000 to 1000 would be well within reason. Given the weak influence of the prior, the actual data should be the primary driver behind the estimate of the posterior, and if the posterior for all γ_{lpe} functions are estimated with approximately equivalent certainty, then it would be reasonable to assume that this results primarily from the posterior being supported by the data, rather than by the priors in the model.

As for the rest of the model, it follows very similarly to the one described in Chapter 4. For every speaker, indexed by j , their central tendency for /ay/ before each phoneme, p , in each context k was estimated. This by-speaker estimate is represented by μ_{jpk}^s .

$$DOB_{1,2,\dots,n.speaker} \quad (5.24)$$

$$l = DOB_j \quad (5.25)$$

$$\mu_{jpk}^s \sim \mathcal{N}(f_{lpk}, \sigma_k) \quad (5.26)$$

In addition, to the contextual effects of the following segments, I have also modeled the effect of vowel duration on /ay/ height. This effect was treated as being the same in all contexts, but I did allow speakers to differ in the strength of this effect. The community level variable is represented by β_d , and a slope term for each speaker is represented by β_j^s .

$$\beta_d \sim \mathcal{N}(0, 1000) \quad (5.27)$$

$$\beta_j^s \sim \mathcal{N}(\beta_d, \sigma^d) \quad (5.28)$$

Finally, word level random effects were also included, and represented by μ_m^w .

$$\mu_m^w \sim \mathcal{N}(0, \sigma^w) \quad (5.29)$$

At the raw data layer of the model, normalized F1 is the outcome variable being modeled, represented by y_i . The duration variable is passed to the model as x_i , and is, in fact $\log(\text{duration}) - \text{median}(\log(\text{duration}))$. The remaining variables are indices for indexing the speaker level and word level effects.¹¹

$$y_{1,2,\dots,n} \quad (5.30)$$

$$J_{1,2,\dots,n} \quad (5.31)$$

$$P_{1,2,\dots,n} \quad (5.32)$$

$$K_{1,2,\dots,n} \quad (5.33)$$

$$W_{1,2,\dots,n} \quad (5.34)$$

$$x_{1,2,\dots,n} \quad (5.35)$$

$$j = J_i \quad (5.36)$$

$$p = P_i \quad (5.37)$$

$$k = K_i \quad (5.38)$$

$$m = W_i \quad (5.39)$$

Each observation is modeled as being drawn from a normal distribution with a speaker specific variance σ_j^s . The mean of this normal distribution is the sum of the speaker level estimate for F1 for the specific following segment and context, μ_{jpk}^s , the word level effect, μ_m^w , and the speaker level duration effect, $\beta_j^s x_i$

$$y_i \sim \mathcal{N}(\mu_{jpk}^s + \mu_m^w + \beta_j^s x_i, \sigma_j^s) \quad (5.40)$$

This model was implemented in Stan, and was set to run with four chains with a 3,500 iteration burn in, and a 3,500 iteration sample. All location parameters which were not defined in the model

¹¹ J is for speaker indices, P is for the following segment (/t/ or /d/), K is for the context (surface or flap), and W is for word indices.

above were given a prior of $\mathcal{N}(0, 1000)$, and all scale parameters were given a prior of $\mathcal{U}(0, 100)$, which for the scale of this data are relatively uninformative priors. The model converged to very stable estimates, based on the Gelman-Rubin Potential Scale Reduction Factor, \hat{R} . Figure 5.10 plots a histogram of \hat{R} for all parameters in the model. For all parameters, \hat{R} is very close to 1.

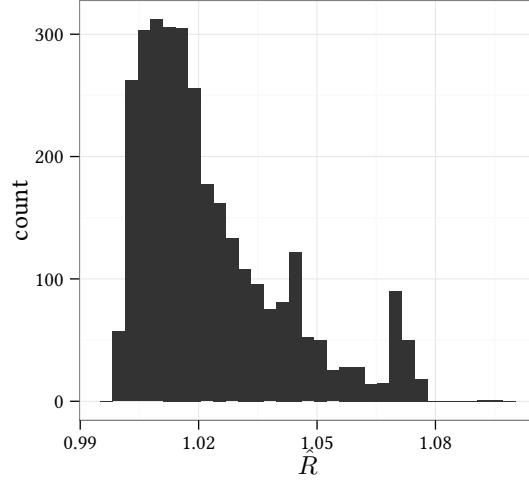


Figure 5.10: \hat{R} for all parameters in the model.

Figure 5.11 plots the estimated F1 trajectories for /ay/ in each context, along with 95% highest density posterior intervals. As it was in Chapter 4, since this model is both Bayesian and fitting non-linear curves, I don't have p-values to report. Rather, there is a 95% probability that the true value lies within the colored band representing the HPD. There is a strong qualitative similarity in the trajectories in Figure 5.11 to those in Figure 5.9, which along with the \hat{R} values close to 1 suggests that the model as described above is an adequate one for the data. Figure 5.12 plots the same estimated F1 trajectories from Figure 5.11, but this time facetting by the following underlying stop in order to foreground the effect of flapping. While /ay/ before /t/ looks nearly identical whether or not that /t/ is flapped, there is a much larger difference for /ay/ followed by flapped and surface /d/. The effect on flapping /t/ and /d/ are highlighted in Figure 5.13, which plots the difference in the curves in each facet of Figure 5.12.

Figure 5.13 plots the difference in height of /ay/ between flaps and surface realizations for /t/ and /d/. Values below 0 mean that flaps are lower than surface realizations, and values above 0

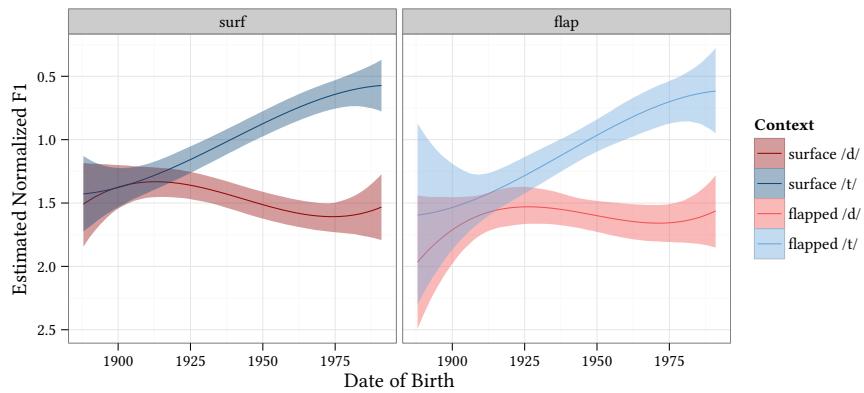


Figure 5.11: Model estimates of /ay/ F1, faceted by surface vs. flapped realizations.

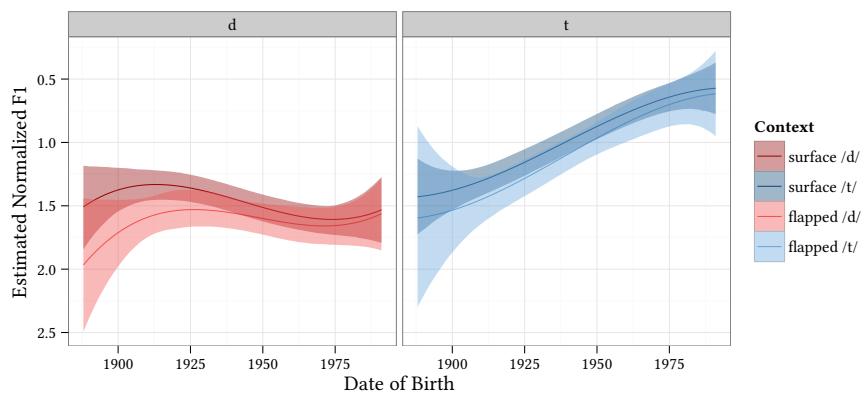


Figure 5.12: Model estimates of /ay/ F1, faceted by /t/ vs /d/.

mean that flaps are higher than surface realizations. Looking at /t/ first, the 95% HPD contains 0 throughout the entire change, meaning that the height of /ay/ is not reliably different between surface /t/ and flapped /t/. The edge of the 95% HPD comes very close to excluding 0 around 1910, but this is also true for /d/, with flapped /d/ being lower than surface /d/. If anything, there appears to be some kind of flapping main effect with /ay/ before flaps being somewhat lower, although not reliably. The lack of any difference between /ay/ before surface /t/ and flapped /t/ is anomalous under the glide peripheralization precursor hypothesis, which would predict that /ay/ before flapped /t/ should pattern more or less like /ay/ before surface /d/ at the onset of the change. The main effect of /ay/ before flaps being lower than before surface realizations is anomalous under the duration precursor hypothesis, which predicted that /ay/ before both /t/ and /d/ flaps should undergo raising. The effect of flapping on /ay/ height, therefore, is not consistent with either of the phonetic precursor hypotheses.

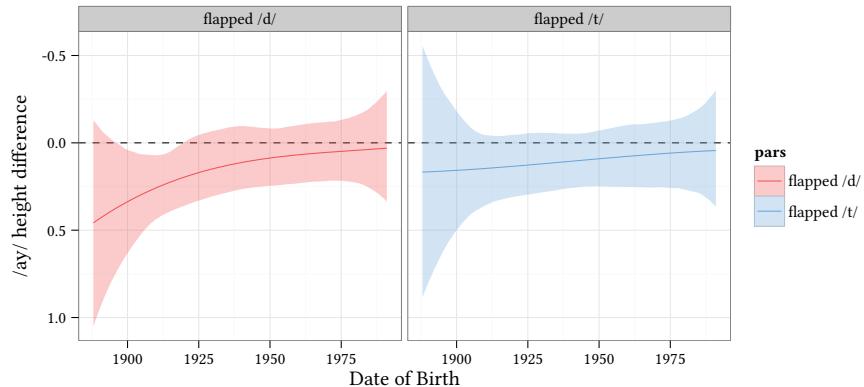


Figure 5.13: The difference in normalized F1 for /ay/ before flapped /t/ and /d/ from surface /t/ and /d/. The y-axis can be understood as $(ayC_{flap} - ayC_{surf})$

Figure 5.14 plots the second relevant comparison, the effect of /t/ within surface realizations and flaps. The way to interpret the “surface” facet of 5.14 is that it plots the height difference in /ay/ before surface /t/ and /d/. The “flap” facet plots the height difference in /ay/ before flapped /t/ and /d/. The effect of following /t/ is virtually identical for flaps and surface realizations. They both begin to exclude 0 at approximately the same time around 1920. Figure 5.15 plots the difference between the surface /t/ effect and flapped /t/ effect. This difference between /t/

effects contains 0 throughout the 20th century, meaning that the difference in /ay/ height between flapped /t/ and /d/ has always been the same as the difference between surface /t/ and /d/.

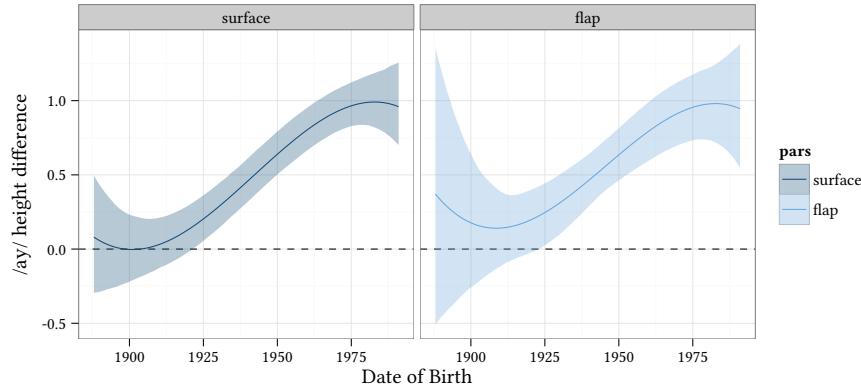


Figure 5.14: The effect of following phonological voice on /ay/ across context. The y-axis can be understood as (ayd-ayt).

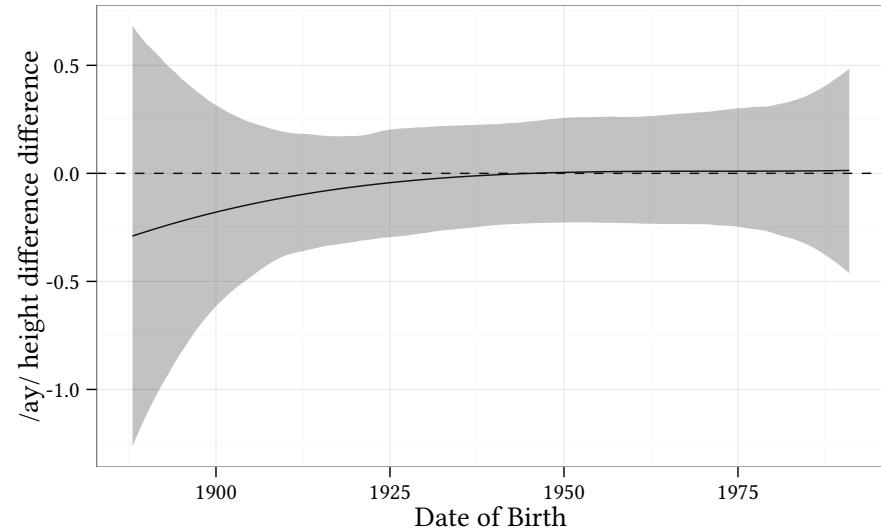


Figure 5.15: The difference in the effect of voicing between surface and flap contexts. The y-axis can be understood as $(ayd_{surf}-ayt_{surf})-(ayd_{flap}-ayt_{flap})$

/ay/ Conclusions and Discussion

The results laid out above are strikingly at odds with a model of conditioned sound change which is based on the accumulation of phonetically conditioned production and perception errors *a la*

Ohala (1981). Neither the rate of change nor the degree to which one context or the other favors the change appears to be proportional to either of the proposed phonetic precursors. Rather, it seems clear that the conditioning of /ay/ raising must make reference to *phonologically* defined categories rather than phonetically defined ones. This phonological differentiation of raised and low /ay/ suggests that a grammatical process like (5.41) entered the phonology at the onset of the change, and was always in an opaque relationship with respect to flapping.

(5.41) ay → –low/___–voice

In many respects, this early introduction of a grammatical process is very similar to the Competing Grammars view of language change from historical syntax beginning with Kroch (1989). Kroch and students have by and large found that syntactic change does not begin in one context and then spread by analogy to others. Rather, changes begin in all contexts at the same time, although some may be more favoring than others and boost the overall rate. Given these results from pre-voiceless /ay/ raising, it is obvious that it did not begin in the most phonetically favoring environment and then analogically spread to other contexts.

One important difference between this case and most syntactic changes is that the introduction of the new phonological process to raise pre-voiceless /ay/ must have been *rapid*, in fact, too rapid to be detectable to the analysis methods I employ in this dissertation. In syntactic change, the change we observe is the rate of use of the new grammatical process, where in this case use of the new phonological process must have reached categorical use nearly immediately, and the change we observe is shifting phonetic implementation of the output of that phonological process. See Chapter 2 §2.3 for the quantitative arguments that this is the case. In Chapter 6, I'll discuss the possibility that there was a precursor phonological process for pre-voiceless /ay/ raising that may solve this rapidity problem for /ay/ raising, but this rapidity is really a problem for the phonological conditioning on sound change discussed in chapter 4 (specifically the effect of /l/ on /ow/ and /uw/), as well as for /ey/, to be discussed next.

5.1.2 /ey/ Raising

The conditioned raising of /ey/ in “checked” position was initially described as a new and vigorous change in Philadelphia in the 1970s (Labov, 2001). However, not as much work has been done on this change as has been done on /ay/, and since conditioned raising of /ey/ is not a feature shared by other dialects (or at least not reported to), a bit more exploratory description is necessary before digging into its conditioning.

To begin with, I have excluded the days of the week (*Sunday*, *Monday*. etc.) because of the relatively frequent lexical variation in these items between /-deɪ/ and /-di:/ which constitutes a cross-cutting factor that is not of particular relevance to the patterns discussed below. I’ve also excluded all cases of /ey/ followed by /g/. There has been a long lasting tendency in Philadelphian phonology to lax /iy/ and /ey/ before /g/ to /ɪ/ and /ɛ/, leading to such shibboleths as *Iggles* (*Eagles*, the local football team) and *beggle* (*bagle*). Figure 5.16 plots ellipses representing the distribution of /ey/ when followed by various stops. The distribution followed by /g/ is clearly outlying, and again, a cross-cutting factor not relevant to the problem at hand.

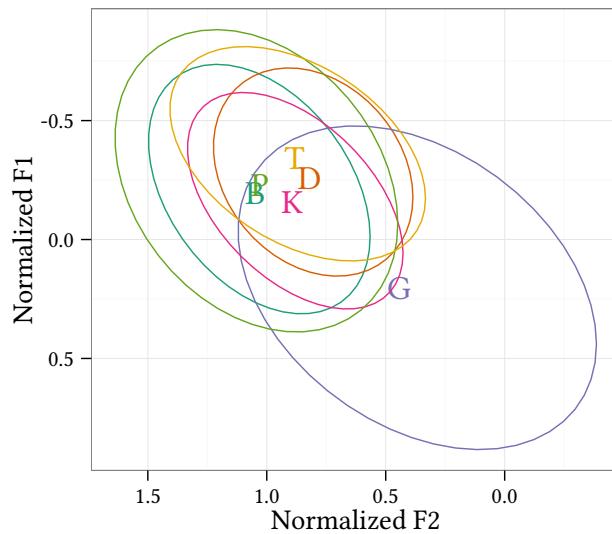


Figure 5.16: Distribution of /ey/ means by following stop.

As for the remaining possible within-word conditioning effects, Figure 5.17 plots most of them out, faceting by voicing and manner, color representing place of articulation. There are no

standout effects, except for that of following /l/, which is more or less flat. For the rest of this section, I'll be collapsing across following consonant, separating out only /l/.

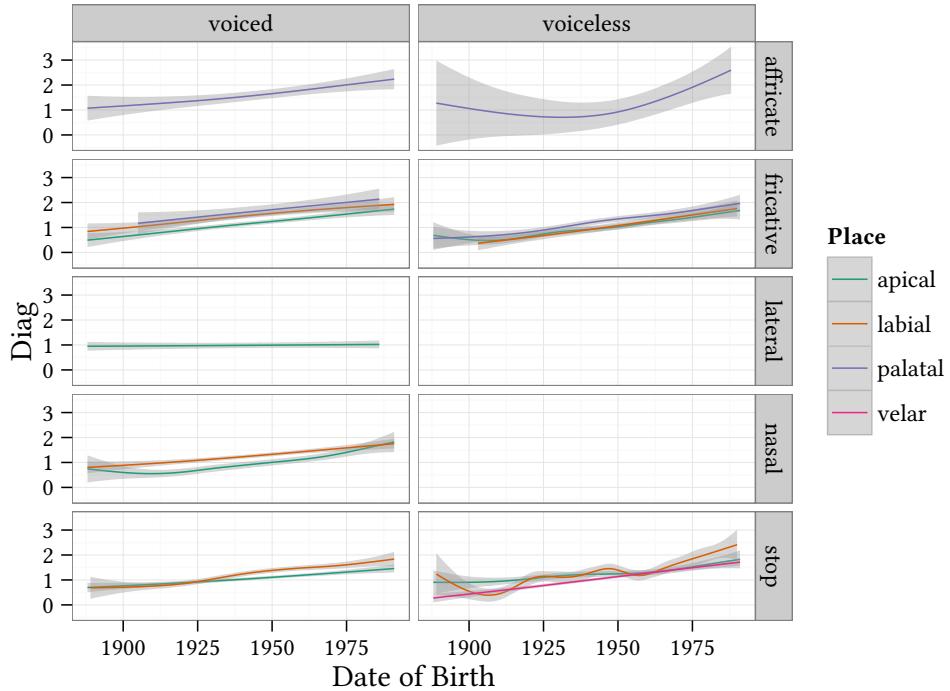


Figure 5.17: The effect of following context on word internal /ey/ raising.

One major question regarding /ey/ raising is how it interacts with syllable structure. It has been mostly defined in the previous literature as raising in “checked” position, while remaining low in “free” position. This distinction has been largely operationalized as being word final (free) versus all other contexts. However, it has not been established whether the distinction between open versus closed syllables plays any role. To determine whether or not syllabic structure plays a role in conditioning /ey/ raising, I wrote a simple syllabifier to categorize the consonants following word internal /ey/ as to whether they were in the onset of the following syllable (making /ey/ the nucleus of an open syllable) or in the coda of the syllable with /ey/ (making it closed). The syllabifier operated over the CMU dictionary style transcriptions for each word, and maximized onset (Kahn, 1976) modulo English phonotactics. There was not enough data of /ey/ followed by /l/ to further subdivide it by syllacticity. Of course, many tokens of /ey/ were pre-hiatus (e.g.

mayor, saying), and this context was also separated out. For now, I've set aside word-final /ey/.

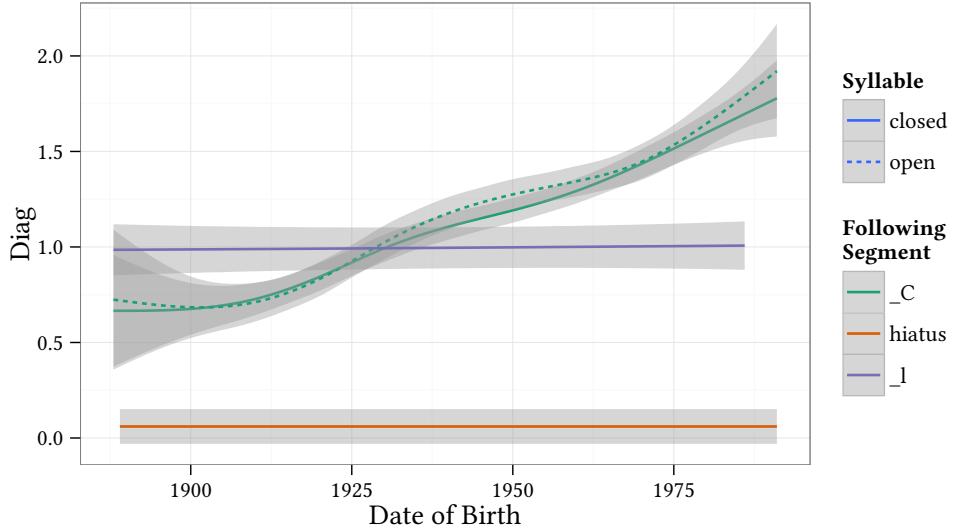


Figure 5.18: Trajectory of word internal /ey/

Figure 5.18 plots cubic regression splines over speaker means. Colors indicate the category of following segment (consonant, hiatus, /l/), and pre-consonantal /ey/ has been further subdivided according to the syllabic relationship between /ey/ and the consonant. There are a few striking results visible in this figure. First, syllabic structure appears to make no difference when the following segment is a consonant. Both *baby* and *babe* undergo the change at the same rate. Second, neither /ey/ followed by /l/ nor /ey/ followed by a vowel appear to undergo the change at all. Third, the set of contexts where /ey/ undergoes the change is not related to the degree to which those contexts appeared to favor change at the beginning. This is a change along the front diagonal of the vowel space, and in the period around 1900, the context where /ey/ was most advanced along the front diagonal was before /l/. However, the change did not take place before /l/ at all. This appears to be another clear example where the context where the change seemed to be happening first is not where it happened fastest (cf. Bailey, 1973).

The difference along the front diagonal between /ey/ followed by /l/ and /ey/ followed by other consonants is very slight in the early period of the change according to Figure 5.18, so more detailed statistics are necessary to establish its validity. To this end, I fit a mixed effects linear

model using the `lme4` package in R. I created a Decade predictor for the model which is equal to $(\text{DOB}-1900)/10$. This allows us to interpret the intercept effects as representing differences between the contexts in 1900, and to interpret the slope effects as representing the degree of change per decade. The fixed effects included in the model were Decade, the following segment (consonant, vowel, /l/), and their interaction. The random effects included random intercepts for speaker and word, as well as a random slope of following segment by speaker, and random slope of decade by word.¹² The model estimates are displayed in Table 5.5 along with t-values. For the purposes of this dissertation, t-values greater than 2 will be taken to indicate a reliable effect.¹³ Since all of the t-values in Table 5.5 are greater than 2, we'll take the reported effects to be reliable.

	C Effects		Interactions		
Intercept	0.63	t=15.14	0.31	t= 2.66	/l/
			-0.55	t=-4.29	V
Decade	0.12	t=14.9	-0.1	t=-4.18	/l/
			-0.12	t=-5.22	V

Table 5.5: Regression Estimates for word internal /ey/ raising. Reference levels: Decade=1900; FolSeg=C. Model formula: $\text{Diag} \sim \text{Decade} * \text{FolSeg} + (\text{FolSeg} | \text{Speaker}) + (\text{Decade} | \text{Word})$.

The intercept for /ey/ followed by consonants is 0.63, with a decade-over-decade rate of change of 0.12. There are approximately 9 decades between 1900 and the most recent date of birth, 1991, meaning that /ey/ rose 1.09 units along the front diagonal, from 0.63 to 1.73. The effect of a following /l/ in 1900 was 0.31, meaning it would take /eyC/ $0.31/0.12=2.55$ decades to reach that level. This effect of /l/ is reliable and substantial in the direction of the change, but /ey/ does not undergo the change in this context. The slope of /ey/ followed by /l/ is estimated to be 0.1 less than /ey/ followed by other consonants, or $0.12 - 0.1 = 0.02$. Given the current specification of the model, it's not possible to determine whether 0.02 is reliably different from 0, but it is unlikely to be. To be sure, I refit the model changing the reference level from a following C to following /l/. The results are displayed in Table 5.6, and, in fact, the t-value corresponding to the slope of

¹²Due to the structure of the data, it is not possible to include random slopes for following segment by word, or decade by speaker.

¹³p-values are not included because they are non-trivial to calculate for mixed effects models (Bates, 2006) and estimation of p-values by MCMC is not yet implemented for models with random slopes.

/ey/ before /l/ is less than 2, meaning that it is not reliably different from 0. The interaction effect of following V is also not reliably different from 0 in Table 5.6, meaning /ey/ followed by /l/ and /ey/ followed by vowels are not reliably different from parallel.

	/l/ Effects		Interactions		
	Intercept	0.94 t=8.12	-0.31 t=-2.66	C	
			-0.87 t=-5.15	V	
Decade	0.02 t=1.08	0.1	t= 4.18	C	
		-0.02 t=-0.78		V	

Table 5.6: Regression Estimates for word internal /ey/ raising. Reference levels: Decade=1900; FolSeg=/l/. Model formula: $\text{Diag} \sim \text{Decade} * \text{FolSeg} + (\text{FolSeg} | \text{Speaker}) + (\text{Decade} | \text{Word})$.

The qualitative impressions from Figure 5.18 are therefore quantitatively supported. Even though /eyl/ started out higher and fronter than /eyC/, it did not undergo the raising and fronting change. In fact, around approximately 1925, /eyC/ crossed over and passed /eyl/.

Unlike /ay/ raising, there has been less work on establishing what the phonetic precursors of /ey/ raising might have been. Even without that background research, however, I believe it can still be established that the context for /ey/ raising is phonologically, not phonetically defined. Strictly phonetic effects, like those discussed in Chapter 4, specifically §4.1, can be seen to be operating early on in /ey/ just before the change begins around the turn of the century. When followed by a vowel, /ey/ is especially low, and when followed by an /l/, it is slightly higher than when followed by other consonants. Whatever phonetic precursor for raising /ey/ may be proposed, it is clear that /eyl/ has more of it, because it starts off in a more advanced position. The fact that /eyC/ comes from behind and overtakes /eyl/ means that the change must have been conditioned by something other than phonetic favorability.

The proposal that I'll put forward here about /ey/ raising is that it is conditioned by following consonants. I'll call the phonological allophone [+peripheral], following the fact that it fronts and raises along the front peripheral track (Labov, 1994). Moreover, the allophone which does not undergo raising and fronting remains with phonetics roughly similar to the realization of /ey/ in Southern English, which Labov et al. (2006, ch 3) argues is [-peripheral].

(5.42) ey → +peripheral /__C

Of course, the first issue that arises is why a following /l/ does not count as a consonant, but there is precedent for an analysis of /l/ as not being a consonant in Philadelphia. Philadelphia has fairly aggressive /l/ vocalization, taking place not only word finally and in codas, but also intervocally and in initial clusters (Ash, 1982). /l/ also has a number of vowel-like effects on preceding vowels in Philadelphia, especially to the long back-upgliding vowels /uw, ow, aw/ as discussed in Chapter 4. Notably, it triggers glide deletion in /aw/, leading Dinkin (2011a) to propose that /l/ is actually the glide in /awl/. Looking only at word internal /ey/, it is not possible to determine whether /ey/ raising is conditioned or not by other glides and liquids, but this will be addressed shortly.

Placing /ey/ raising in the grammar.

The fact that /ey/ raising excludes the most phonetically favoring context is fairly good evidence for its phonological conditioning. It is also possible to see how the proposed phonological process interacts with other phonological processes, like I did for /ay/ raising. Unfortunately, it appears that /ey/ raising applies transparently with respect to affixing, and may even apply at the phrase level.

The easiest way to demonstrate that /ey/ raising applies transparently with respect to affixing is to plot the trajectories if the relatively high frequency words *day* and *days*. Figure 5.19 plots speaker means and cubic regression splines for these two words. *Days* appears to behave like every other /eyC/ context by undergoing the change, while *day* remains low.

Figure 5.20 plots an extended comparison like the one in Figure 5.19. It compares /ey/ followed by inflectional /-z/ and /-d/ to uninflected forms of the same words where /ey/ is word final, as well as to /ey/ followed by /z/ and /d/ which are not exponents of any morpheme. /ey/ appears to pattern the same way regardless of whether the following consonant is inflectional morphology or not, and in all cases the uninflected form remains low. Again, this qualitative impression from Figure 5.20 is supported quantitatively in Tables 5.7 and 5.8, where the main effects and interactions for the comparison consonants (/z/ and /d/ which are not exponents of agreement)

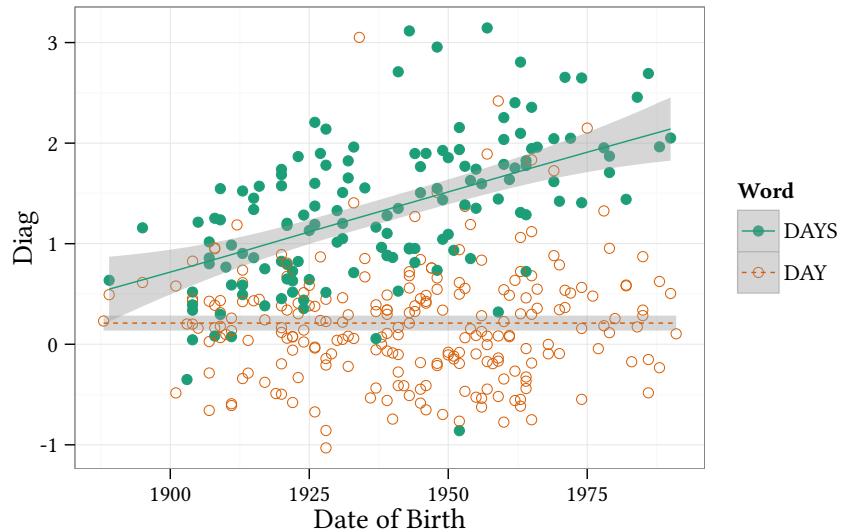


Figure 5.19: Trajectories of the words *day* and *days*.

are not reliably different from the inflectional morphemes.

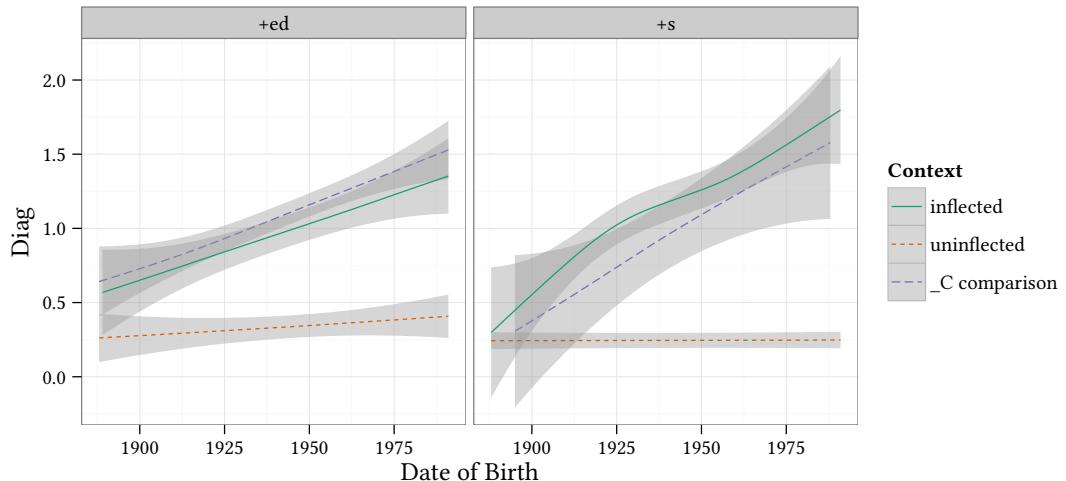


Figure 5.20: Comparison of word final /ey/ to inflected versions of the same words, as well as /ey/ followed by non-morphological /z/ and /d/.

Additionally, there appears to be some conditioning of /ey/ raising for word final /ey/ by the onset of the following word. Figure 5.21 plots the trajectories of word final /ey/ divided up

	+ed Effects		Interactions		
Intercept	0.49	t=4.17	0.24	t= 1.67	non-inflectional C
			-0.31	t=-2.14	uninflected
Decade	0.09	t=3.3	-0.01	t=-0.25	non-inflectional C
			-0.08	t=-2.15	uninflected

Table 5.7: Regression Estimates /ey/ raising in inflected vs uninflected words. Reference levels: Decade=1900; Context=/+ed/. Model formula: $\text{Diag} \sim \text{Decade} * \text{Context} + (\text{Context}|\text{Speaker}) + (\text{Decade}|\text{Word})$.

	+s Effects		Interactions		
Intercept	0.48	t= 4.95	-0.34	t=-1.56	non-inflectional C
			-0.26	t=-2.34	uninflected
Decade	0.14	t=6.48	0.05	t= 1.07	non-inflectional C
			-0.12	t=-5.11	uninflected

Table 5.8: Regression Estimates /ey/ raising in inflected vs uninflected words. Reference levels: Decade=1900; Context=/+s/. Model formula: $\text{Diag} \sim \text{Decade} * \text{Context} + (\text{Context}|\text{Speaker}) + (\text{Decade}|\text{Word})$.

by the onset of the following word (the “etc.” category consists of following pauses, coughs, laughs, etc.). The same conditioning effects appear to be in place, except within a substantially compressed range and a shallower slope (Figure 5.21 has the same y-axis range as Figure 5.18 for this comparison).

Both the fact that /ey/ raising interacts transparently with inflection and that it applies across word boundaries indicate that this must be a low level phonological process. In a stratal approach to phonology like Lexical Phonology or Stratal OT, this process would be taking place postlexically, or at the phrase level. I’ll revise the process in (5.42) to reflect this.

$$(5.43) \text{ ey} \rightarrow +\text{peripheral} / __C]_{\text{phrase level}}$$

Now, there is frequently ambiguity between phrase level or postlexical rules and strictly phonetic processes, but the case for /ey/ raising being a phonological process is still evident at the phrasal level because of the exclusion of /l/ from the conditioning environments. There is more diversity in the range of contexts following /ey/ when it is word final than when word internal,

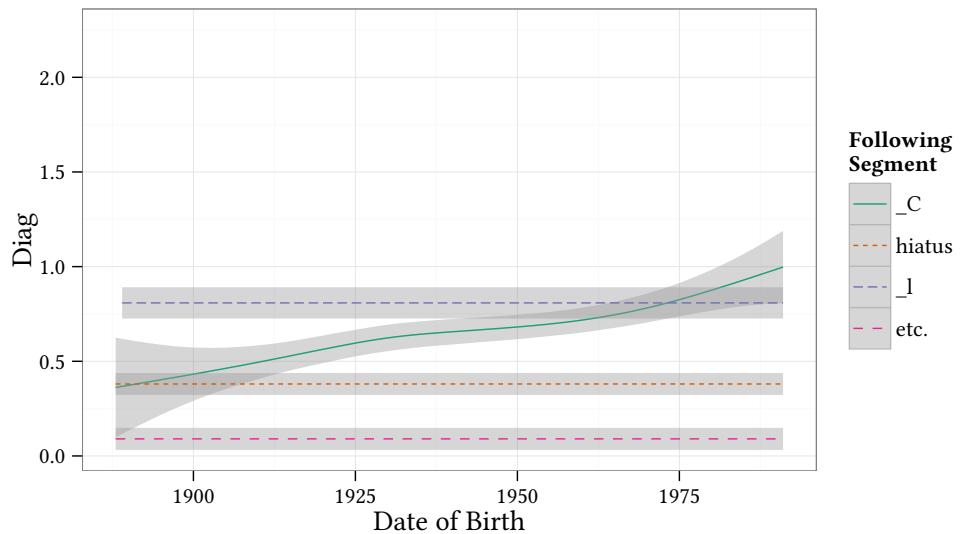


Figure 5.21: Trajectory of word final /ey/

meaning it's possible to compare /ey#C/, /ey#l/, and /ey/ followed by other liquids and glides like /r, w, y/. Figure 5.22 does exactly this, plotting cubic regression splines over speakers' means for /ey/ followed by various contexts. The erratic behavior of /ey#r/ is almost certainly due to its small volume of data, yet it does generally follow the trend of becoming higher and frontier. /ey#y/ and /ey#w/ have essentially identical trajectories as /ey#C/. The only non-participating contexts are /ey#l/, /ey#V/ and /ey#/.

The analysis that /ey/ raising is phonologically conditioned by following consonants appears to be supported again by the analysis of word final /ey/. It appears to be phonological for two reasons. First, the exclusion of /ey#l/ from the set of undergoers cannot seem to be done on a phonetic basis, as other phonetically similar segments (/r, w/) do undergo the change. Second, the following contexts appear to arrange themselves into two categorically separate classes: triggers and non-triggers.

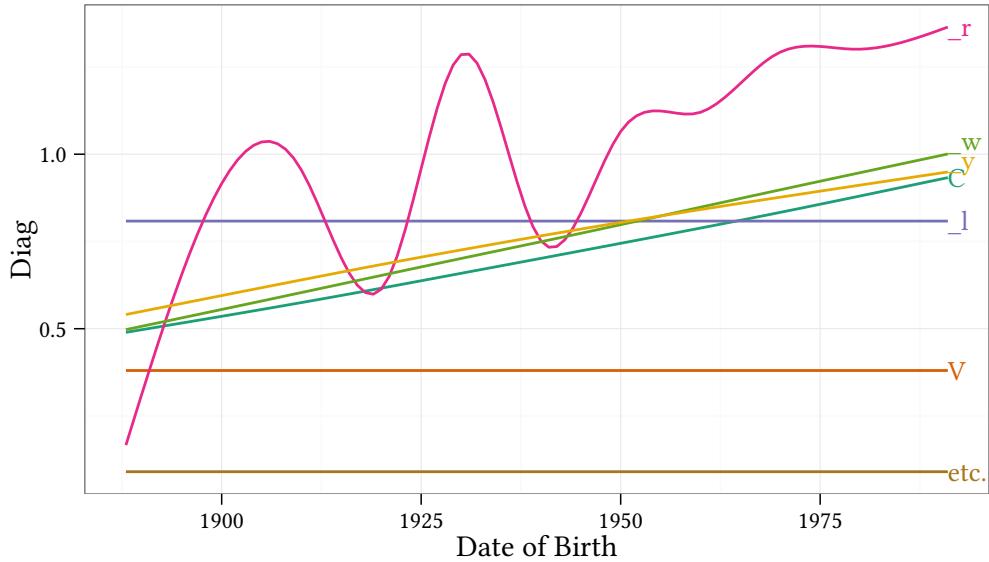


Figure 5.22: Trajectory of word final /ey/

5.1.3 /ay/ and /ey/ summary

In this section, I have examined divergent diachronic trajectories that occurred within the phonemic categories /ay/ and /ey/ and determined that they are best explained by appealing to phonological, rather than phonetic conditioning factors. In neither case did the change occur in proportion to the phonetic favorability of the contexts where it could have. In the case of /ay/ raising, we should expect /ay/ raising before flaps to have either patterned with surface /t/ or surface /d/ on phonetic grounds, depending on the precursor theory we care to adopt. However, /ay/ raising appeared to pattern strictly according to the underlying phonological voice of the following segment. One may try to argue that this result could be achieved through some kind of analogical process without resorting to phonological processes, so that [ai] in *write* analogizes to *writer*. Of course, whatever explanation based on analogy one comes up with must also allow for raised /ey/ in *days* not to analogize to *day*.

In the case of /ey/ raising, the exclusion of /l/ from the set of environments where the change took place is anomalous on a number of grounds. The exclusion of /eyl/ from undergoing the

change must be made on grounds other than its phonetic properties (due to the fact that phonetically similar segments /r/ and /w/ *did* undergo the change) or its phonetic favorability for the change (due to the fact that before the change began, it appeared to be the *most* favoring environment under consideration here). Since there is independent evidence that /l/ is not quite consonantal in Philadelphia, and since /ey/ did not raise before vowels, I'll argue that the definition of the conditioning environment for /ey/ raising is that the following segment must be a consonant phonologically.

The conditioning of these two changes support my proposal that the target of phonetic changes are surface phonological representations. For both of these changes, I have posited phonological processes that separate the phonemic categories into two phonological allophones, only one of which undergoes the change. In the next section, I'll demonstrate that phonetic changes frequently target phonological natural classes as a whole, and argue that this is the result of changing phonetic implementation of the phonological features which define these natural classes.

5.2 Natural Class Patterns

The most commonly discussed multi-vowel shifts are chain shifts, a useful typology of which is provided by Labov (1994). The observed patterns in chain shifts are frequently described in terms of maximizing the “margin of security” between vowels (Martinet, 1952), or by the maximal dispersion of vowel contrast (Liljencrants and Lindblom, 1972; Labov, 1994, 2001; Flemming, 2004 *inter alia*). For example, the low-back merger of the LOT and THOUGHT vowels (/ɑ/ and /ɔ:/, or /o/ and /oh/ in Labovian class labels) is implicated in at least two kinds of chain shifts. The first is the Canadian Shift¹⁴ (Clarke et al., 1995; Labov et al., 2006), where the phonetic gap created by the merger of /o/ to a backer and higher position is filled by short-a, /æ/. The remaining front short vowels then lower, due to the phonetic gap created by /æ/ retraction.

The second kind of shift implicated in the low-back merger involves the lowering of /ʌ/ into the gap left behind by /o/, which has been attested in Pittsburgh (Labov et al., 2006). Figure 5.24 is a schematic diagram of this Pittsburgh Chain Shift, and Figure 5.25 plots the vowel system of a

¹⁴also known as the California Shift or Third Dialect Shift

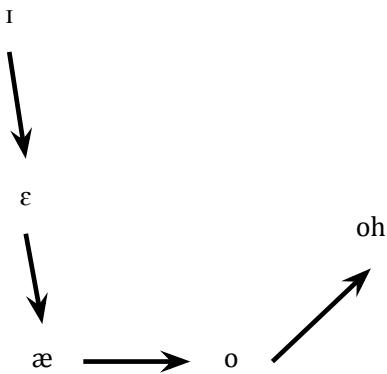


Figure 5.23: The Canadian Chain Shift in reaction to the merger of /o/ and /oh/. Modified from Clarke et al. (1995).

representative speaker from the Atlas of North American English dataset.

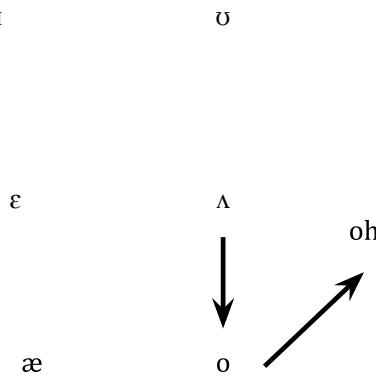


Figure 5.24: The Pittsburgh Shift in reaction to the merger of /o/ and /oh/. Modified from Labov et al. (2006).

The retraction of /æ/ and the lowering of /ʌ/ into the gap created by the merger of /o/ and /oh/ can be motivated on largely phonetic grounds, as can the subsequent lowering of /ɛ/ and /ɪ/ in the Canadian Shift. Labov (2001) offers one such explanation for chain shifting whereby the creation of a phonetic gap allows for outliers to be included in the estimation of the phonetic target for that category. In the case of the low-back merger, before /o/ merges into /oh/, it has a phonetic realization of approximately [ɑ]. Outlying pronunciations of /æ/, due to production

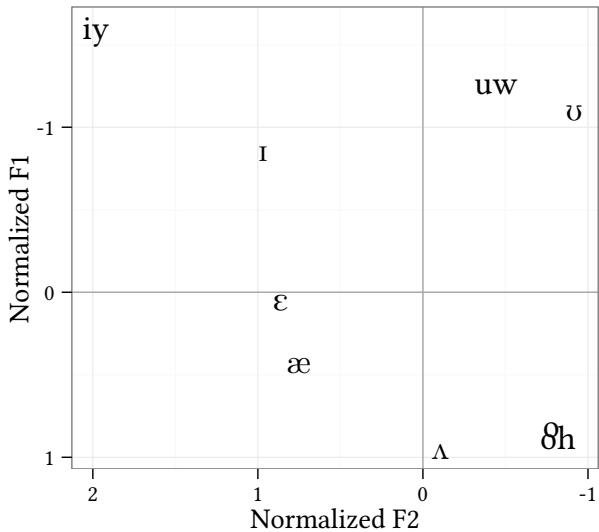


Figure 5.25: Vowel system of a 66 year old woman from Pittsburgh.

error or any other cause, which are produced sufficiently back enough may be misinterpreted as /o/. These exceptionally back realizations of /æ/ wind up not being factored into the estimate for the phonetic target of /æ/, since they were miscategorized into /o/. However, once /o/ merges into /o/, moving from a phonetic realization of [ɑ] to [ɒ: ~ ɔ:], there is no longer a phonemic category with a realization of [ɑ]. These exceptionally back pronunciations of /æ/ are now less likely to be miscategorized into /o/, meaning that now they will be included in the estimation of the target for /æ/, leading to /æ/ retraction. For the Canadian Chain Shift, this process repeats with /ɛ/ leading to /ɛ/ lowering. For The Pittsburgh Shift, this process played out for /ʌ/ instead of /æ/.

Other implementations of this basic model of chain shifting exist, from more teleological ones, like Martinet's (1952) margins of security or Flemming's MINDIST constraints on contrast , to more mechanistic ones, like those relying on exemplar models and agent based models (de Boer, 2000, 2001; Ettlinger, 2007, e.g.), or Boersma and Hamann's (2008) model of bi-directional cue constraints . A more vexing phenomenon for all of these models which adequately explain either /æ/-retraction or /ʌ/-lowering in reaction to the low-back merger is when *both* happen. The original description of the shift in Clarke et al. (1995) included both /æ/-retraction and /ʌ/ lowering

(shown schematically in Figure 5.26), and an ANAE speaker from Winnipeg who appears to be exhibiting both is shown in Figure 5.27.

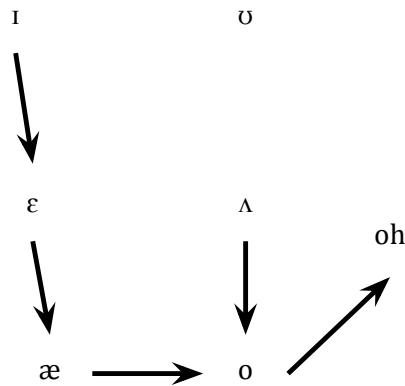


Figure 5.26: The Canadian Chain Shift with /Α/ lowering. Modified from Clarke et al. (1995).

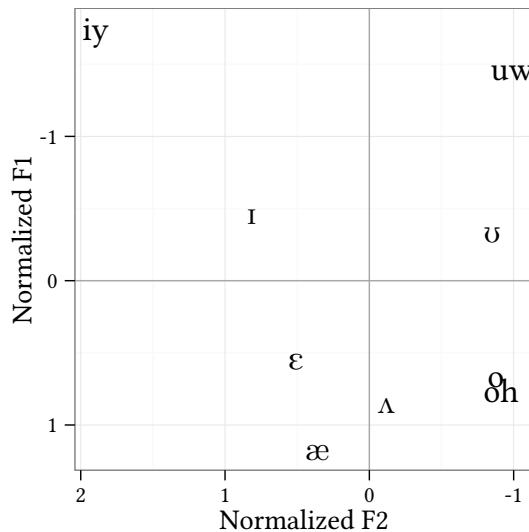


Figure 5.27: Vowel system of a 36 year old woman from Winnipeg.

The focus of this section, however, is a third reported pattern of the Canadian Shift which is also inexplicable in terms of gap creation and filling, the maximization of contrast, or any other operationalizations of those concepts. Boberg (2005) reports that in Montreal, the dominant pattern of movement for /ε/ and /ɪ/ is their parallel *retraction*, rather than a rotation in a chain shift. Figure 5.28 plots the normalized F1 and F2 values for the 3 generational groups Boberg

(2005) studied.¹⁵ The parallel retraction of /æ/, /ɛ/, and /ɪ/ has also been reported as occurring in Columbus, OH by Durian (2012, ch 5), who found relatively strong and significant correlation in the mean F2 of /æ/ and /ɛ/ across speakers. In these cases, the only phonetic pressure explicable in terms of gap creation and filling is for the retraction of /æ/ into the space vacated by /o/. The subsequent parallel retraction of /ɛ/ and /ɪ/ cannot be similarly explained, especially since this parallel retraction compresses the distance between the short front vowels and the short back vowels in a way unexpected under a maximal dispersion kind of account. Instead, it appears that in Montreal and Columbus, the retraction of /æ/ is generalizing to the other short front vowels along phonological dimensions.

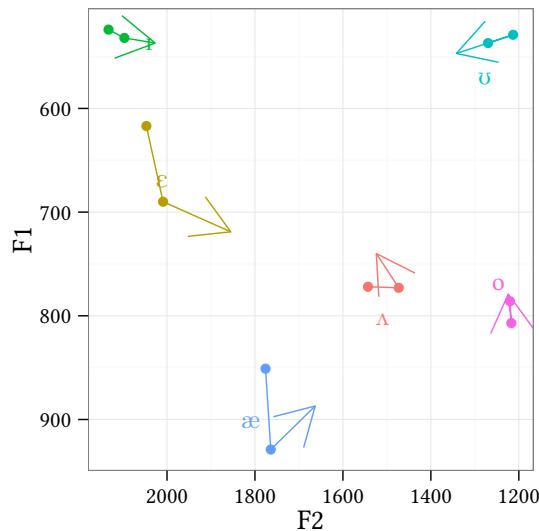


Figure 5.28: The Canadian Parallel Shift. Data from Boberg (2005) Table 4. Points represent 3 generational groups. F1 and F2 have been normalized according to the method reported in Labov et al. (2006)

5.2.1 Back vowel fronting in Philadelphia

Philadelphia has so far been resistant to the low-back merger (Labov et al., 2006), and there is no evidence in the PNC of any retraction of the short front vowels. However, the fronting of the back

¹⁵Boberg (2005) employed a one-factor Nearey log-mean normalization, similar to that used in the Atlas of North American English (Labov et al., 2006). The generational groups were defined as “(1) born before 1946, (2) born 1946-1965, and (3) born after 1965.”

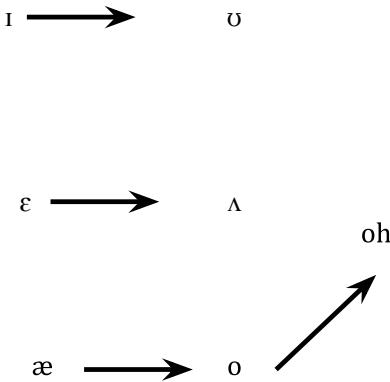


Figure 5.29: The Canadian Parallel Shift. Modified from Boberg (2005).

up-gliding vowels /aw/, /ow/, /uw/ in Philadelphia appear to progress in parallel, a noted trend in many North American Dialects (Labov et al., 2006; Fridland, 2001; Baranowski, 2008; Durian, 2012, *inter alia*). The only plausible “triggering event” for back vowel fronting is proposed by Labov (2010a). He argues that the merger of post-coronal /iw/ and /uw/ (e.g. *dew* [dju:] and *do* [du:]) triggered the eventual fronting of /uw/. As discussed in Chapter 4, the rate of change data for [Tuw] is ambiguous with respect to whether [Tuw] and [uw] are phonologically distinct, but in combination with the plausible historical event of [Tiw ~ Tuw] merger, and the results in Labov, Rosenfelder, and Fruehwald (2013) which found that [Tuw] patterns separately from [uw] along social dimensions, I believe it is reasonable to conclude that [Tuw] and [uw] are phonologically distinct. Of particular interest in this chapter, however, is the degree to which fronting occurs in parallel between all of these back upgliding vowels, so I will be excluding [Tuw] from the data in this section, as well as /aw/, /ow/ and /uw/ followed by /l/, which I found to be phonologically distinct allophones in Chapter 4. Figure 5.30 plots the trajectory for these back vowels along normalized F2. /aw/ starts out much fronter than /ow/ and /uw/, and this is probably because its nucleus has a different phonological specification for backness (in fact Dinkin (2011a) argues that /aw/ has merged with /æl/). Strikingly, however, all three vowels appear to front at nearly the same rate, reach a maximum at about the same time, and then begin to reverse together.

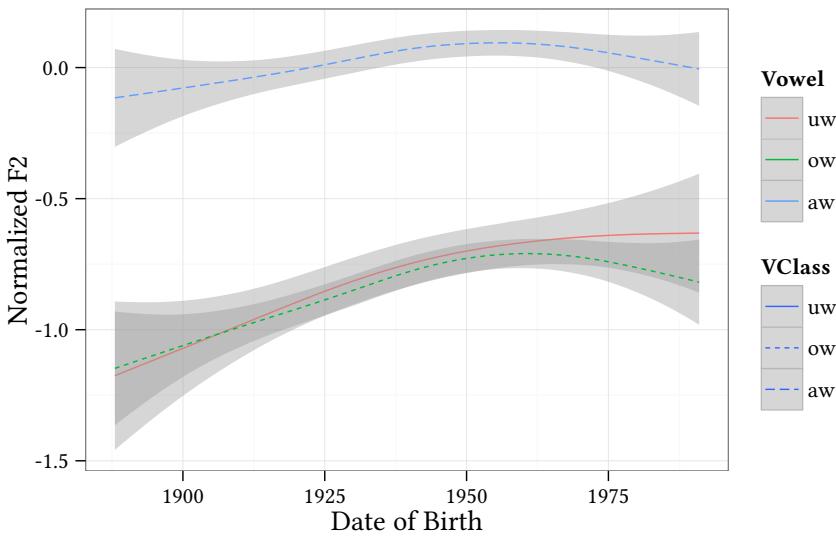


Figure 5.30: The fronting of back vowels in Philadelphia

If we plot speakers' means for /uw/, /ow/ and /aw/ against each other, as Figures 5.31, 5.32 and 5.33 do, we can see a relatively strong relationship between the frontness of one vowel and the others, meaning speakers who have very fronted /aw/ are likely to also have very fronted /ow/ and /uw/. This relationship is strongest for the vowels which are adjacent in height (/aw/ and /ow/, and /ow/ and /uw/), and weakest for the vowels which are two steps away from each other in height (/aw/ and /uw/). However, correlation tests find that the correlation between all three pairwise comparisons of /aw/, /ow/ and /uw/ are significant.

Table 5.9 displays the results of statistical tests using three different correlation statistics. The well known Pearson's r is a measure of the linear correlation of the two vowels, and Spearman's ρ is a similar statistic which measures the correlation of the two vowels given any monotonic function. Kendall's τ is a measure of the concordance of two vowels. For example, taking two speakers and their /ow/ and /aw/ measurements, if the first speakers' /ow/ and /ow/ were *both* fronter than the second speakers', this would be a "concordant" pair of speakers. On the other hand, if only the first speaker's /ow/ were fronter than the second speaker's, but their /aw/ was backer, this would be a "discordant" pair of speakers. Kendall's τ is the proportion of all pairwise

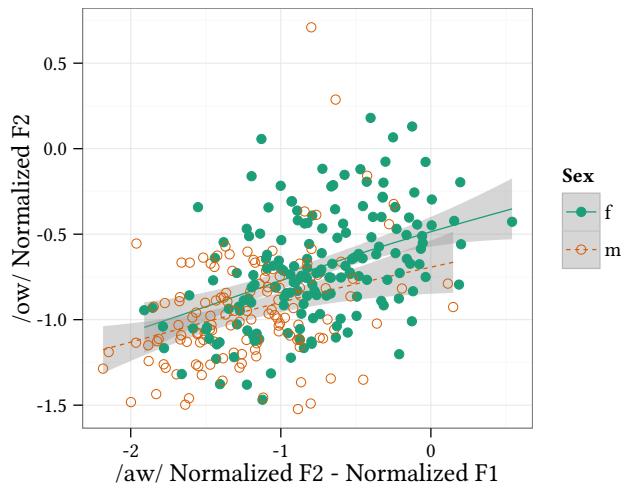


Figure 5.31: The relationship between /aw/ and /ow/ across speakers.

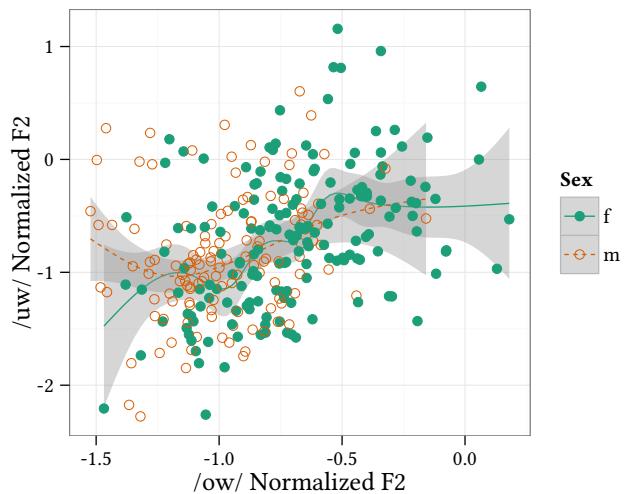


Figure 5.32: The relationship between /uw/ and /ow/ across speakers.

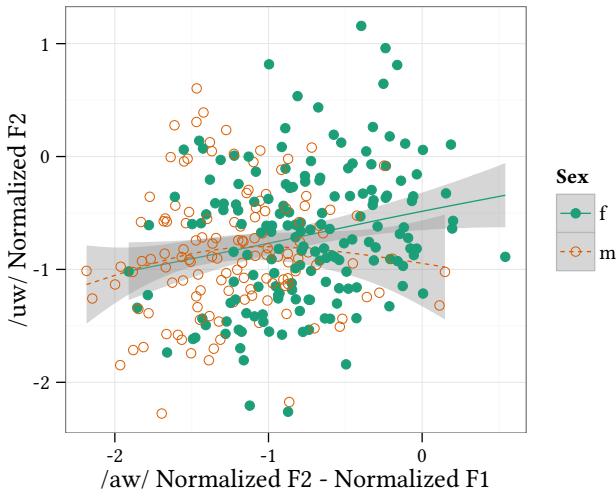


Figure 5.33: The relationship between /uw/ and /aw/ across speakers.

comparisons of speakers which are concordant. For all three of these correlation statistics, the relationship between /aw/, /ow/ and /uw/ are significantly and positively correlated.

	Pearson's r	Spearman's ρ	Kendall's τ
/uw/~/ow/	0.38 ($p < 0.001$)	0.4 ($p < 0.001$)	0.28 ($p < 0.001$)
/ow/~/aw/	0.5 ($p < 0.001$)	0.52 ($p < 0.001$)	0.36 ($p < 0.001$)
/uw/~/aw/	0.22 ($p < 0.001$)	0.2 ($p = 0.001$)	0.14 ($p = 0.002$)

Table 5.9: Correlation of /uw/, /ow/ and /aw/ across speakers (reported p-values have been Bonferroni corrected within each test statistic).

My argument is that this pattern of parallel fronting and retraction of /aw/, /ow/ and /uw/ is due to their shared membership in a phonological natural class. I've already argued in Chapter 4 that these three vowels are phonologically active together with respect to a following /l/. The particular label that should be used to define the set {aw, ow, uw} is not of great importance. If Mielke (2008), is correct, the set of phonological features may in fact be “emergent,” meaning language learners identify sets of phonologically active segments, and assign a label to them. Following the Labovian approach to the phonology of English vowels (Labov et al., 1972; Labov, 2006b; Labov et al., 2006), I'll use the label “+Vw” for this set with the understanding that this is

the label for a feature. In Philadelphia at least, [+Vw] both defines a set of phonologically active vowels, but is also an input to phonetic implementation. At the beginning of the change, the phonetic implementation rule would have mapped the nuclei of [+Vw] vowels to a relatively back target. Then the rule changed, and began mapping the nuclei of [+Vw] vowels to slightly frontier targets, dragging all of the [+Vw] vowels forward. This change continued until approximately 1950, at which point it reached a maximum and began to reverse, dragging all of the [+Vw] vowels back.

A reasonable counter proposal to the phonological one I've put forward here is that the correlation of frontness of /aw/, /ow/ and /uw/ is not due to a shared phonological feature, thus shared phonetic implementation, but is rather due to a shared sociolinguistic evaluation. This is, in fact, the argument Watt (2000) makes for the parallel behavior of /ey/ and /ow/ in Tyneside English, which I will discuss more below. Labov, Rosenfelder, and Fruehwald (2013) argue that the reversal of /aw/, /ow/ and /uw/ fronting is due to Philadelphia's dialectal reorientation from being a Southern dialect city to a Northern one. It could be possible that these three vowels pattern together not because they are phonologically related, but because fronted pronunciations for them are all understood as being "Southern." Of course, it is possible for the reversal of fronting to be socially motivated, but it does not eliminate the need to appeal to their shared phonological features to explain their parallel behavior. As Labov (2006b, [1966]) pointed out, social evaluation cannot be tied to strictly phonetic categories. In New York City, [iə] was a stigmatized realization of /æh/, and it was frequently corrected to [æ] or [a]. However, this negative evaluation was restricted to the phonological-phonetic mapping of /æh/ to [iə], since the mapping of /i:r/ to [iə] due to NYC r-lessness did not result in the correction of *beer* to [bæ:]. In the case of [+Vw] fronting, the social evaluation of fronted /aw/, /ow/ and /uw/ as being "Southern" explains the motivation for reversing their fronting trend, but it does not explain why this reversal generalized to all three vowels immediately, nor why it did not effect any other vowel. For example, the lax, pre-vocalic allophone of /ey/ which was discussed above is similar to the realization of this vowel in the Southern Vowel Shift (Labov et al., 2006). The fact that this allophone of /ey/ remained low, and did not begin to raise at the same time that [+Vw] vowels began to retract is unaccounted

for under the strictly social evaluation analysis. The combination of a negative social evaluation of Southern speech in combination with a phonological natural class [+Vw] can explain both the reversal of [+Vw] fronting *and* the fact that it applied to all and only [+Vw] vowels.

However, it is possible to try to factor out the effect of social evaluation of [+Vw] fronting and see if the frontness of /aw/, /ow/ and /uw/ is still correlated across speakers. Figure 5.34 plots cubic regression splines which were fit using generalized additive mixed effects models. For each vowel, for each gender and educational level, I fit a gamm where the outcome variable¹⁶ was predicted by a cubic regression spline of date of birth. Also included in the model were random intercepts by speaker and by word. I'll be using the by-speaker random intercepts as a sort of by-speaker residual, by which I mean that any by-speaker effects which are not accounted for by their date of birth, sex, and educational level should be captured by their random intercept. As Figures 5.35, 5.36, and 5.37 show, even after controlling for social factors as much as possible, there is still a fairly strong relationship in frontness across these three vowels. That is, speakers with exceptionally front /aw/ for their gender, educational level and birth cohort are also likely to have exceptionally front /ow/ and /uw/. Due to the constraints on how random intercepts are estimated, it does not seem appropriate to do significance testing for their correlation. However, I still calculated the correlation statistics, and compare them to the correlation statistics estimated just over speaker means in Table 5.10. In some cases, the correlation of the random intercepts is weaker, but over all they appear to be largely similar.

	Pearson's r		Spearman's ρ		Kendall's τ	
	means	random effects	means	random effects	means	random effects
/uw/~/ow/	0.38	0.38	0.4	0.39	0.28	0.26
/ow/~/aw/	0.5	0.34	0.52	0.34	0.36	0.23
/uw/~/aw/	0.22	0.14	0.2	0.14	0.14	0.1

Table 5.10: Comparison of correlation statistics based on speaker means to those based on random effects.

Another alternative for the parallel [+Vw] fronting is proposed by Durian (2012), who suggests it can be understood as “phonetic analogy.” According to Durian (2012), “phonetic analogy refers

¹⁶The outcome variable was normalized F2 for /ow/ and /uw/, and normalized F2 minus normalized F1 for /aw/.

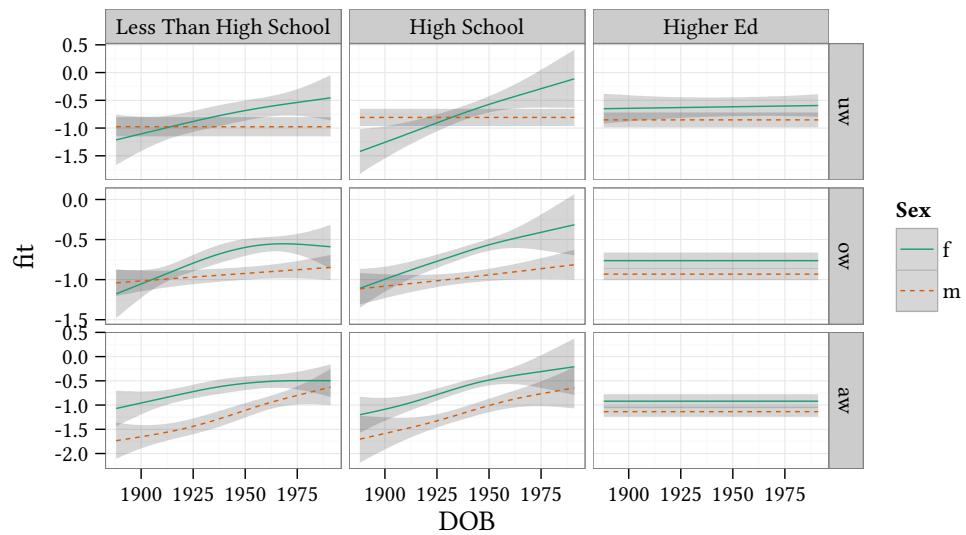


Figure 5.34: Cubic regression spline fits from the generalized additive mixed effects models.

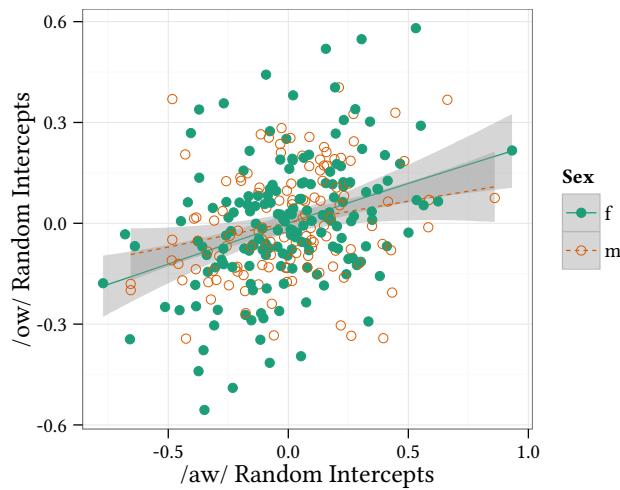


Figure 5.35: The relationship between /aw/ and /ow/ across speakers.

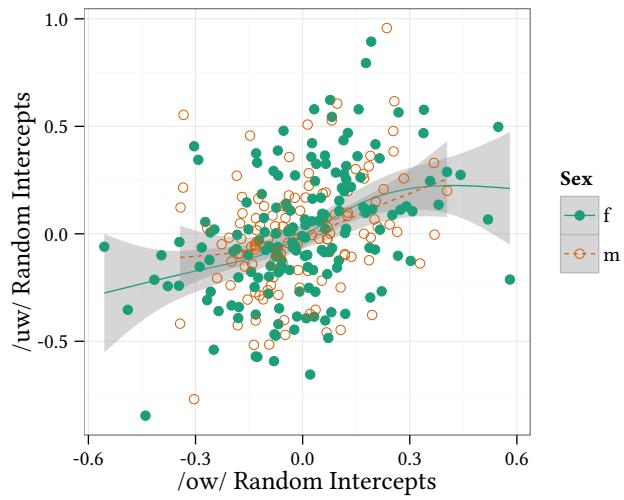


Figure 5.36: The relationship between /uw/ and /ow/ across speakers.

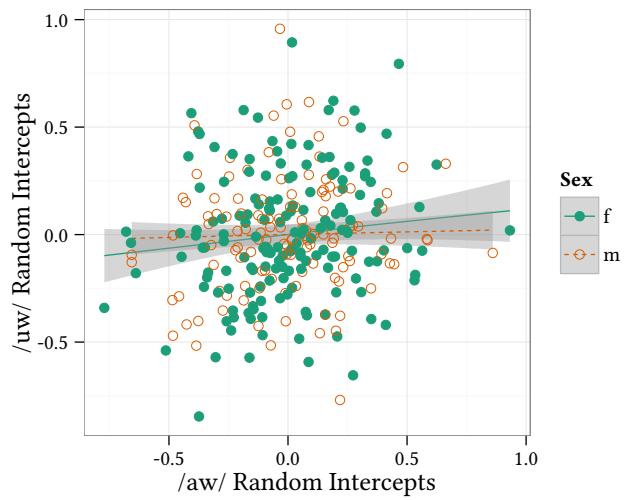


Figure 5.37: The relationship between /uw/ and /aw/ across speakers.

to a process in which a likening between two entities, along some definable dimension, in this case phonetic (and/or possibly phonological) is drawn between two entities, by a speaker.” The substantive difference between Durian’s proposal and the one put forward here is how we propose the fronting trend generalizes to all [+Vw] vowels. Durian cites the fact that the conditioning effects of preceding coronals are so similar between /uw/ and /ow/ as evidence that fronting is generalizing from *context to context*. The degree of fronting of /uw/ when preceded by /s/ analogizes from /suw/ to /sow/, and so on, essentially turning the parallel fronting of /uw/ and /ow/ into multiple simultaneous changes, one for each phonetic context. My proposal is that there is only one change, and it is to the phonetic implementation of [+Vw], which drags along all [+Vw] vowels with it. I attribute the fact that the conditioning is so similar across phonetic contexts to the uniform and independent set of coarticulatory processes in the language. If the effect of preceding coronals or following fricatives is strictly due to natural acoustic or coarticulatory phenomena, then they *should* be consistent across phonetic context. Ultimately, the appeal to phonetic analogy and my appeal to phonetic implementation of phonological features will be able to account for most of the same phenomena. I believe that my approach will prove to be more fruitful, however. It is more restrictive, meaning it predicts fewer kinds of possible sound changes than phonetic analogy, and it provides a linking hypothesis between observed phonetic changes and phonological representation, meaning that it can lend support to other sorts of phonological investigation. For example, the fact that {aw, ow, uw} are a phonologically active set of vowels (losing their glides when followed by /l/) and the fact that they all front in parallel are intrinsically related, under my account, to the fact that they have a shared phonological feature which I am calling [+Vw].

5.2.2 Long-ingliding vowel lowering in Philadelphia

Another parallel vowel shift in Philadelphia is the lowering of /oh/ and /æh/. Like the broader Mid-Atlantic region, Philadelphia’s realization of /oh/ is a mid tense ingliding vowel, ranging somewhere from [ɔ:] to [uə] (Labov, 2001, 2006b; Labov et al., 2006). Also, like New York City, Philadelphia has a split short-a system, with one allophone being tensed and raised under a num-

ber of complex conditions to something between [æə] and [iə]. Both of these vowels are undergoing change in both Philadelphia and New York City. Becker and Wong (2010) report that the complex conditioning of /æh/ is breaking down in favor of something more like a nasal system, and Labov, Rosenfelder, and Fruehwald (2013) found a similar pattern in younger speakers in Philadelphia. Also, Becker (2010) found that /oh/ was lowering among white New Yorkers on the Lower East Side, and Labov, Rosenfelder, and Fruehwald (2013) found a similar pattern obtaining in Philadelphia. For the sake of this discussion, the reorganization of the Philadelphia short-a system to a nasal system is an orthogonal issue. I am strictly interested in whether the tense and ingliding short-a is moving in parallel with tense and ingliding /oh/, so the tokens of short-a under consideration here are those which would be tense under either phonological system, meaning only short-a which appear before front nasals /m, n/ in closed syllables. Figure 5.38 plots the relationship between /oh/ and /æh/ height, and there is a substantial correlation, similar in magnitude to what Labov (2006b, [1966], p.349, Fig 14.1) found in New York City. Using the same correlation statistics as we did for [+Vw] fronting above, we can see similarly strong and statistically significant correlations for [+Vh] height in Table 5.11.

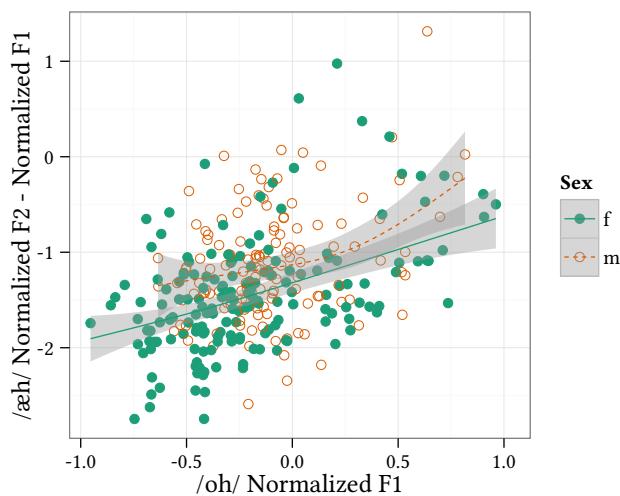


Figure 5.38: The correlation of /oh/ and /æh/ in Philadelphia

Again, the argument could be made that the correlation between /æh/ and /oh/ could be due to social rather than phonological factors. Both raised /æh/ and /oh/ are subject to negative social

	Pearson's r	Spearman's ρ	Kendall's τ
/uw/~/ow/	0.43 ($p < 0.001$)	0.39 ($p < 0.001$)	0.27 ($p < 0.001$)

Table 5.11: Correlation of /oh/ and /æh/ across speakers.

evaluation (Labov, 2001), and Labov, Rosenfelder, and Fruehwald (2013) found that both education and gender had an effect on the height of both of these vowels. Just like I did for [+Vw] vowels, I fit a generalized additive mixed effects model for each vowel for each gender and educational level where the outcome variable was predicted by a cubic regression spline over date of birth. Random intercepts for speaker and word were also included, and the speakers' random intercepts taken to indicate the individual level variation which was not explained by the other effects. Figure 5.39 plots the by speaker random intercepts for /oh/ and /æh/ against each other. The same basic correlation still exists after accounting for social factors, as the correlation statistics for the random intercepts show in Table 5.12.

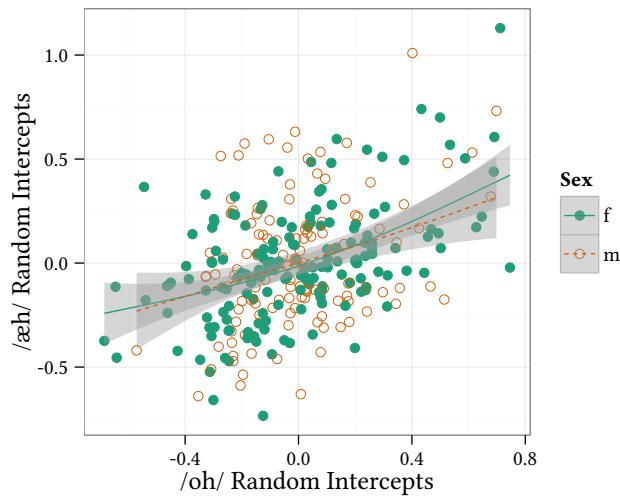


Figure 5.39: /oh/ and /æh/ Random Intercepts

	Pearson's r		Spearman's ρ		Kendall's τ	
	means	random effects	means	random effects	means	random effects
/oh/~/æh/	0.43	0.43	0.39	0.36	0.27	0.25

Table 5.12: Comparison of correlation statistics based on speaker means to those based on random effects for /oh/ and /æh/.

5.2.3 Searching for more parallel shifts.

Parallel shifts seem to suffer from having been under-labeled. The conceptual notion of “chain shifts” is relatively salient in the field, and subsequently many sound changes are described in terms of chain shifting. Parallel shifts, on the other hand, are much less commonly described, but this may not be due to their rarity in reality. Durian (2012) points out, for example, that the relatively obvious case of parallel [+Vw] fronting, has occasionally been described as being a chain shift, even though its mechanics must be different. If the concept of parallel shifting becomes more salient in the field, my belief is that more sound changes will be described as such. Labov (2010a, ch 5) addresses the question of whether all chain shifts could be recast as parallel shifts of a sort. The last two stages of the Southern Vowel Shift, for example, is the lowering of /ey/ and /iy/ to a non-peripheral position, a process which could be described in my framework as a change in the phonetic implementation rule for non-low front upgliding vowels.

$$(5.44) \quad \begin{bmatrix} \text{Vy} \\ \text{-low} \end{bmatrix} \rightsquigarrow 0.5 \text{ Diag}$$

However, Labov (2010a) stresses the importance of “bends in the chain of causality,” where the initial trigger for the subsequent changes cannot be subsumed into the same generalization which describes them. Keeping with then Southern Vowel Shift Example, the triggering change is argued to be the monophthongization of /ay/ (Labov et al., 2006). Whether /ay/ monophthongization is a phonological or phonetic process, it cannot be subsumed under the generalization in (5.44). The same point can be made about the low-back merger of /ɑ/ and /ɔ/ triggering the Canadian Shift. So while the notion of sequential, chained, movement is applicable to the triggering of many shifts, a large portion of the subsequent shifts can be recast as parallel. In this section, I’ll briefly

discuss some examples of phonetic changes which I believe could be parallel shifts, even though they have not been described as being so.

[+Vw] glide fronting in Southern British English

In their description of koine formation in the town of Milton Keynes, Kerswill and Williams (2000) provide the following transcriptions as possible innovative variants of the [+Vw] vowels.

(5.45) /uw/ [ɥ:], [y:]

/ow/ [yɑ], [ɪə]

/aw/ [æʊ], [a:ə], [ɛ:], [ɛʏ], [εɪ]

Of particular interest to me here is the possibility of fronting the glide in /ow/ and /aw/. Kerswill and Williams (2000) don't describe how common the fronted glide is for /aw/, nor how the frontness of the /aw/ glide might covary with the frontness of the /ow/ glide, but glide fronting is a striking possibility especially when compared to North American [+Vw] fronting. For the most part, when the nuclei of [+Vw] vowels front in North America, the glide retains its high back target.¹⁷ It would seem that in Southern British English, the phonetic implementation of [+Vw] which is changing not only affects the nucleus, but the glide as well. The stand out exception would appear at first to be /uw/, which Kerswill and Williams (2000) describes as fronting as a long monophthong. However, Chládková and Hamann (2011) provide an acoustic analysis of Southern British English /uw/ fronting and find that, in fact, fronted /uw/ is diphthongal with a backward trajectory.

There are two important facts here. First, the same sort of phonetic change (fronting) is affecting the same sets of vowels, but result in two different kinds of phonetic outcomes highlights the non-triviality of phonetic change. Presumably, North American and Southern British English speakers have comparable articulatory and perceptual systems such that the fact that North American [+Vw] fronting retains a back glide target and SBE [+Vw] fronting includes the glide cannot be explained on grounds of such naturalness. The second important fact, however, is that

¹⁷In Philadelphia, the glide may lower to [j], but it certainly doesn't front.

even though North American and Southern British English differ from each other, they are internally consistent. If in North America, the glide for /ow/ fronted while the glides for /uw/ and /aw/ remained back, or vice versa for Southern Britain, that would be inexplicable under my account of changing phonetic implementation rules. The fact that the fronting is consistent within each country suggests that within each country there is not three separate fronting processes, but rather *one*. In North America there is one change occurring to the phonetic implementation rule for [+Vw] which leads to a fronter nucleus, and all [+Vw] vowels are affected. In Southern British English, there is a different change occurring to the phonetic implementation rule for [+Vw] which leads to fronted nuclei and glides, and all [+Vw] vowels are affected.

GOAT and FACE diphthongization in the North of England.

Haddican et al. (forthcoming) report on a change in /ow/ and /ey/ in York, UK, where they are diphthongizing from traditional [e:] and [o:] to [ei] and [ou], respectively. Figure 5.40 is from Haddican et al. (forthcoming), and plots the relationship across speakers in the diphthongality of /ey/ and /ow/, measured using the Euclidian distance between vowel onset and 90% into the vowel. The correlation in these two measures is much stronger than the ones presented already in this chapter. Haddican et al. (forthcoming) report a significant Spearman's ρ of 0.9.

Watt (2000) reports a similar kind of trend in Tyneside English which is complicated by the fact that ingliding variants of /ey/ and /ow/ ([ɪə] and [ʊə], respectively) are possible. Watt (2000) finds that use of ingliding variants tracks understood patterns of covert prestige, and is used the most by working class men. However, the overall rate of ingliding variants decreases with age, and the rate of diphthongal variants is increasing. The rate of use of each kind of variant is tightly correlated across speaker groups and stylistic contexts. Watt (2000) rejects the kind of internally motivated parallelism that I am advocating here, largely because his data does not support the historical development of these vowels that many people have argued for, namely that given in (5.46).

$$(5.46) \quad \begin{array}{ccc} e: & \rightarrow & i\theta \\ o: & \rightarrow & u\theta \end{array} \quad \begin{array}{ccc} & \rightarrow & ei \\ & & ou \end{array}$$

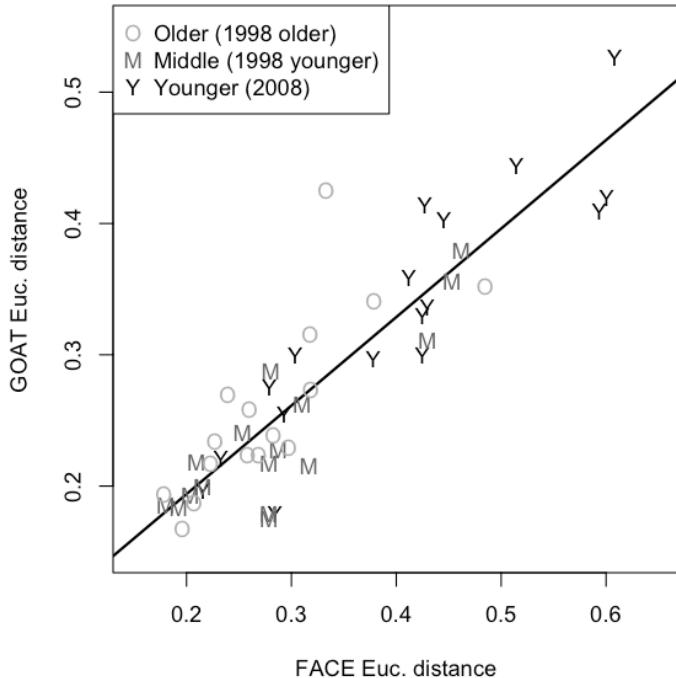


Figure 5.40: The relationship between /ow/ and /ey/ diphthongization. Figure from Haddican et al. (forthcoming, Figure 2)

Rather, he argues that the introduction of the diphthongal [eɪ] and [oʊ] is introduced through a process of dialect leveling. As I argued for the case of [+Vw] fronting above, however, while the motivation for a particular sound change may be socially defined, we must also explain which set of sounds the change applies to, and which are excluded in some way, and in this case, as it was for [+Vw], there is an obvious phonological dimension at work. The combination of a social evaluation with a phonological dimension across which it applies has the greatest explanatory adequacy.

The parallel diphthongization of /ey/ and /ow/ in Northern British English is a nice example of a phonetic change which is not just a movement along the front/back or high/low dimension. Here, what is changing is the phonetic realization of the long mid vowels, or perhaps just the phonetic realization of their second mora, depending on how we want to represent diphthongs phonologically.

5.2.4 Parallel Shifts are Changing Phonetic Implementations of Phonological Features, but there are Complications

As I've been making the argument that parallel phonetic changes are the result of changing phonetic implementations of phonological features, I have not tried to provide a formalization of this process. There are a few reasons for this. To begin with, the data I am working with are formant measurements of vowel nuclei, but it is not clear whether the mental representations of phonetic form are the same as these formant measurements. Redoing this study with articulatory measurements would probably not resolve the issue (although it would certainly reveal other interesting properties (Mielke et al., forthcoming)), since it is not altogether clear that the mental representation of phonetic forms is strictly articulatory either. As Pierrehumbert (1990) pointed out, “[p]honetic representation is one of the most difficult problems in linguistics,” and I will not be attempting to resolve that problem here. Instead, I'll refer to these changes as occurring along abstract phonetic dimensions, like “backness” or “diphthongality,” which can be reinterpreted in acoustic or articulatory terms as necessary.

Aside from taking the phonetic measurements too literally, a second problem is that a number of the natural classes which are moving in parallel require multiple phonological features to define them under most feature theories. For example, the parallel retraction of the short front vowels in Montreal and Columbus does not include the front long vowels. Taking into account that for the moment we should understand the phonetic dimensions that phonology-phonetic interface rules map to as being more abstract than just literal acoustic or articulatory dimensions, the interface rule at the beginning of the phonetic change would have to look something like (5.47), where the possible values for the phonetic dimension of backness are understood as ranging from 0 to 1. The interface rule in (5.48) is a later stage, as the vowels have moved further back.

$$(5.47) \quad \begin{bmatrix} -\text{back} \\ -\text{long} \end{bmatrix} \rightsquigarrow 0.1 \text{ backness}$$

$$(5.48) \quad \begin{bmatrix} -\text{back} \\ -\text{long} \end{bmatrix} \rightsquigarrow 0.3 \text{ backness}$$

It is crucial that the long front vowels be excluded from undergoing this change. One option to avoid using two features to pick out the set of short front vowels for undergoing this change would be to define an ad hoc emergent feature (in the sense of Mielke (2008)) for this class, perhaps [+V̄]. However, if the phonetic frontness of these short front vowels were determined by this [+V̄] feature, then the phonetic frontness of /iy/ and /ey/ would be phonologically accidental, meaning /iy/ and /ey/ would not fall into a phonological natural class with /ɪ/, /ɛ/ and /æ/ with respect to frontness. While this is not necessarily undesirable *a priori*, there is some empirical evidence to suggest that /iy, ɪ/ and /ey, ɛ/ minimally differ. For example, a number of dialects neutralize these vowels before /l/ (Labov et al., 2006), leading to the *feel*~*fill*¹⁸ and *bail*~*bell* merger. If the phonological representation for /iy, ey/ were [–back, +long], and the phonological representation for /ɪ, ɛ/ were [–back, –long], the process of neutralization before /l/ could be understood as (5.51) or (5.50).

$$(5.49) \begin{bmatrix} -\text{back} \\ +\text{long} \end{bmatrix} \rightarrow [-\text{long}] / __l$$

$$(5.50) *[\text{+long}]l$$

However, if the phonetic backness of /ɪ, ɛ/ were not defined by the same phonological feature as it is for /iy, ey/, then to map /iy/→/ɪ/, we would need to eliminate /iy, ey/’s specification for the phonological feature [±back], and replace it with the phonological feature for /ɪ, ɛ/, a decidedly more complex process.

$$(5.51) \begin{bmatrix} -\text{back} \\ +\text{long} \end{bmatrix} \rightarrow \begin{bmatrix} \text{back} \\ \text{long} \\ +\check{V} \end{bmatrix} / __l$$

The phonological inelegance of the above proposal also has empirical problems on the basis of parallel shifts. For example, for the sake of expository clarity, I defined an *ad hoc* phonological feature [+Vw] which defines /uw, ow, aw/ as a natural class. However, in a number of regions of the Midland and South, back vowel fronting *also* affects /ʊ,ʌ/(Labov et al., 2006; Fridland, 2001;

¹⁸<http://val-systems.blogspot.com/2008/04/feel-bag.html>

Labov, 2010a). In Philadelphia, where this is not the case, the back vowel fronting change could be formalized as in (5.52)¹⁹

$$(5.52) \quad \begin{bmatrix} +\text{back} \\ +\text{long} \end{bmatrix} \rightsquigarrow 0.9 \text{ backness} > 0.8 \text{ backness}$$

In the Midland and South, on the other hand, the phonological definition of the vowels undergoing the change is more underspecified, thus pulling in more vowels.

$$(5.53) \quad [+back] \rightsquigarrow 0.9 \text{ backness} > 0.8 \text{ backness}$$

The conclusion to be drawn from this discussion of parallel shifts is that while they present fairly clear evidence that phonological natural classes as defined by phonological features are targets of phonetic change, the implications of this result depends a great deal on the theory of phonological representation and of phonetic implementation one wants to adopt. Minimally, it must be the case that the objects which are the targets of phonetic change must also be the inputs to phonetic implementation. If the process of phonetic implementation is restricted to map just one phonological feature to one phonetic dimension, then a more complex theory of phonological representation and computation is necessary to account for problems like harmonizing the retraction of just the short front vowels with their phonological relationship to the long front vowels. If the phonology phonetics interface can map bundles of features to phonetic dimensions, then the problem becomes one of limiting the power of the interface rules. For example it would be undesirable for $[-\text{back}, -\text{long}]$ to map to a very back target, and have $[-\text{back}, -\text{long}, +\text{low}]$ map to a very front target, because it would eliminate the isomorphism between phonetic quality and phonological relatedness which is both the primary cue for linguists doing phonological analysis, and presumably also language learners.

Proposing a resolution to this problem would be an overreach at the moment, but I hope to have at least demonstrated that studying patterns of language variation and change will prove central to resolving them.

¹⁹I set aside here the extra complication that /aw/ is probably not phonologically [+back] in Philadelphia.

5.3 Conclusions

In this chapter I have reported and analyzed results with two overarching patterns.

- (5.54) Divergent patterns of change within a vowel category which are best attributed to categorical allophones created by the phonology.
- (5.55) Parallel, or convergent patterns of change across vowel categories which are best attributed to phonetic change targeting phonological natural classes.

These examples highlight the key thesis of this dissertation that gradient phonetic changes must be understood in terms of their relationship to categorical phonological representations. As a consequence, a full understanding of these changes can't be obtained without also attempting to understand the system of phonological representation, organization and computation, but at the same time, these changes provide valuable evidence for trying to understand the systems of phonological representation, organization and computation.

Chapter 6

Against Gradual Phonologization

This chapter will serve as a synthesis of the results from Chapters 4 and 5. I will address the challenges these results pose for the most commonly accepted views of sound change, as well as their implications for theories of phonology, phonetics, and language acquisition.

6.1 Conventional Wisdom Regarding Sound Change

There is a conventional wisdom regarding conditioned sound changes like those that I've investigated here that appears to be roughly comparable across disparate research programs that conditioned phonetic changes are the product of gradual accumulation of errors. The results reported in the previous two chapters cast doubt on the gradualness of phonologization. That is, the categorical phonologization of phonetic change appears to occur at the onset of these sound changes, and does so so rapidly that a transition period from pre-phonologization to post-phonologization is not observable. So as to avoid knocking down strawmen, I'll first outline a frequently referenced formulation of this conventional wisdom, then describe how it has appeared in a number of research programs.

I believe the formulation by Ohala (1981) is most representative of the conventional wisdom I'm addressing, even though other researchers depart from this approach either in detail or mechanism. As was mentioned, in Chapter 4, Ohala (1981) proposes a model for back vowel-coronal consonant coarticulation which is based on natural coarticulatory properties. Figure 6.1 presents

Ohala's schematic diagram of this process, whereby the sequence /ut/ in a hypothetical language is coarticulated to a phonetic realization of [yt]. At this historical stage, sound change, understood as change in speakers' linguistic competence, has not yet happened, as listeners are still successfully recovering the surface [yt] production as underlying /ut/. The ontological status

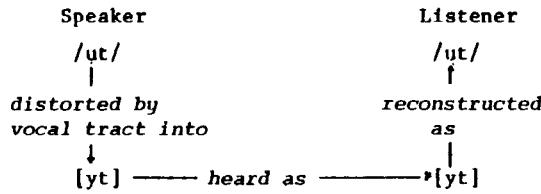


Figure 6.1: Pre-Change Coarticulation, from Ohala (1981)

Ohala assumes for the coarticulation of /ut/ to /yt/ is rather clear from his wording: “distorted by vocal tract.” This is more or less a fact about the contingencies of living in a human body and communicating with a physiological apparatus, rather than speakers’ intention or cognitive system.

Of course, this entire dissertation is devoted to the question of how the observed properties of a language’s sound system ought to be apportioned to different explanatory models, and there is good reason to apply this same kind of reasoning and argumentation when trying to determine whether an effect is due to purely physiological contingencies, or to the language specific system of phonetic alignment and interpolation constraints. For example, in her discussion of vowel duration, Keating (1985) points out that while some people have argued that pre-voiceless vowel shortening has a physiological basis on the grounds that it is a nearly universal effect in the world’s languages, it is, in fact, only *nearly* universal. She found that Polish does not exhibit pre-voiceless vowel shortening at all. Assuming there is nothing physiologically different between speakers of Polish and speakers of other languages, then we must conclude that there is not some proportion of the pre-voiceless shortening effect which is irreducibly physiological. The physiological basis of pre-voiceless vowel shortening, or /ut/ coarticulation, is salvageable if we say that instead of actually *producing* these effects, physiological contingencies *prefer* language specific phonetics which do. Regardless of the exact nature of the physiological or (as Ohala (1981) made sure to emphasize) the perceptual basis of these phonetic effects, the key point is

that they are grounded in properties of the world external to the system of linguistic competence and acquisition. They then percolate up into speakers' linguistic competence through systematic misattribution.

Grounding phonetic changes in the natural systems of production and perception has the benefit of deriving the fact that some kinds of sound change are relatively common, and that they are typically phonetically "natural." Additionally, once these phonetic changes become phonologized and added to the grammar, the explanation for their apparent naturalness can be tied to their origins in phonetic change which in turn have their origins in natural phenomena. As such, it is unnecessary to posit phonetic naturalness constraints on phonological processes, as their observed phonetic naturalness is an historical artifact (Blevins, 2004; Hale and Reiss, 2008).

I would posit, however, that given a hypothetical phonetic change for any speech sound along a single phonetic dimension which is conditioned by one additional factor, that a sufficiently clever analyst could construct a plausible explanation for its naturalness. It appears that for many researchers the naturalness of phonetic change is definitional, rather than a result of empirical investigation. Garrett and Johnson (2011) do point out that the inverse of many common sound changes are unattested. One example they give is that while the palatalization of [k] to [tʃ] before front vowels is common, the backing of [tʃ] to [k] before front vowels is unattested. However, if [tʃi] to [ki] were attested, the explanation for its naturalness is given by Ohala (1981) as hyper-correction, that is, listeners misattributing the phonological target of /tʃi/ as being a coarticulated form of /ki/. Additionally, Kiparsky (2006) provides an elegant counter argument that purely historical accounts of phonetic naturalness alone cannot account for typological gaps. He lays out five hypothetical scenarios where sequences of common and phonetically natural sound changes could produce languages with a productive voicing contrast, but with only voiced word final obstruents, and argues that despite proposals to the contrary, there is no such language attested. A more probabilistic way to phrase Kiparsky's argument is that the rate of attestation of languages with word final voicing (possibly 0) is disproportionately low given the frequency with which sound changes that *could* produce such a pattern happen. I am not arguing here that phonetic changes aren't grounded in natural phenomena, but merely that the sheer obviousness of this

assumption should not be taken for granted.

6.1.1 This conventional wisdom across research programs

I'll briefly outline how this conventional wisdom of error accumulation regarding phonetic change is formulated in a number of research programs here.

Evolutionary Phonology

As outlined by Blevins (2004), the mechanisms of phonetic change assumed by Evolutionary Phonology are very similar to those proposed by Ohala (1981). The three C's of Evolutionary Phonology are CHOICE, CHANCE, and CHANGE, and all three are cases of listeners failing to correctly reconstruct the intentions of speakers. For CHOICE, Blevins (2006) gives the example of a speaker intending to say /tu?ələŋ/, and producing the variants [tu?ələŋ], [tu?ələŋ], [tu?ləŋ]. A listener then chooses one of these variants as the underlying form for the lexical entry, and if that happens to be [tu?ləŋ], then syncope has occurred. This mechanism of CHOICE is not quite adequate in detail to account for the phonetic changes I'm investigating here. In Chapter 2, I argued that phonetic changes don't progress as shifting probabilities over discrete options, but rather as a continuous shift through phonetic space. Moreover, a more realistic formulation of CHOICE would have speakers acquiring probability distributions over the available variants. Nevertheless, CHOICE is the most compatible EP mechanism with the conditioned phonetic changes investigated here, where speakers produce a distribution of phonetic variants, and listeners reconstruct new expectations over those distributions.

It is also worth noting that Blevins' (2004, 2006) formulation of the CHOICE mechanism is also incompatible with my theoretical commitments. Specifically, if the phonetic implementation is qualitatively different from its phonological representation, then it is not possible for the phonetic production of a speaker to be wholesale adopted as an underlying form. Rather, it must be translated into a surface phonological representation by the language specific phonetics, then processed by the phonology. When no distinction is made between surface phonetic production and phonological representation, then it is, in some sense, trivially true that phonological innovation

occurs simultaneously with the onset of phonetic change.

At any rate, the primary driving force behind sound change in the Evolutionary Phonology model is the accumulation of listeners' errors in reconstructing the intentions of speakers.

Exemplar Theory

Exemplar theories run the gamut with regards to the degrees of abstractness they allow. For example, Bybee and McClelland (2005) appear to rule out any abstractness beyond the stored phonetic memory traces when they say that

The innovation in this approach is that language knowledge is not stored in the form of items or rules, but in the form of changes to the strengths of connections among simple processing units.

Pierrehumbert (2006), on the other hand, advocates a more hybridized theory, where phonetic memory traces are associated with phonological categories. This latter position appears to be closer to the mainstream of exemplar theoretic research, so it is this position that I will be referring to when I discuss "Exemplar Theory," although the dynamics of error accumulation are essentially the same under most formulations of ET (e.g. Bybee, 2002).

Simulations of sound change under Exemplar Theory, of which Pierrehumbert (2001) and Garrett and Johnson (2011) are good examples, all involve the same basic mechanism of sampling with replacement. When a speaker has the intention of producing a particular speech segment, they sample from their phonetic memory traces and average over them. Typically, either the sample, the averaging, or both, will be weighted by the individual exemplar's "activation strength," which may be a function many factors including the time since the exemplar was originally perceived, the time since the exemplar was last activated, the exemplar's typicality²⁰, and a number of other potential factors. This average becomes the speakers' new phonetic intention, which they then produce. Of course, production (and perception) is an imperfect process, so the value which gets stored back in the listener's exemplar cloud is perturbed by this systematic error. Figure 6.2 plots

²⁰Garrett and Johnson (2011) implement the down-weighting of atypical exemplars by excluding them from memory upon perception, but this is equivalent to storing them and giving them a 0 activation strength.

an example of simulated phonetic drift from an exemplar theoretic simulation, based on Pierrehumbert (2001).

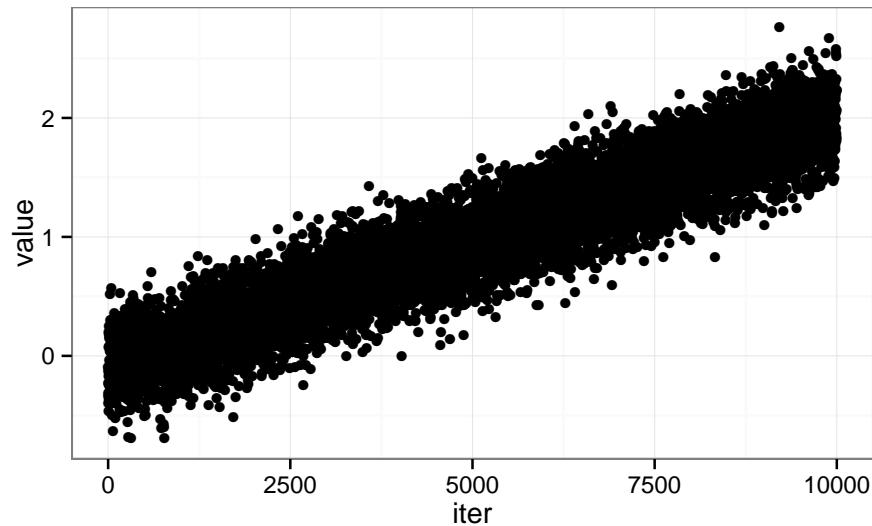


Figure 6.2: Phonetic drift based on exemplar simulation. Model based on the description in Pierrehumbert (2001).

The primary driving force in sound change under Exemplar Theoretic models is the noise introduced by the production-perception feedback loop. When there is a systematic bias to the noise, the exemplar cloud will drift in the direction of that bias. Pierrehumbert (2001) describes this bias in terms of lenition, but coarticulatory drift like that proposed by Ohala (1981) would produce a similar result.

Phonologization

The notion of “phonologization,” whereby a phonetic pattern becomes a phonological one, is central to a number of research programs which posit a qualitative difference between phonetics and phonology, including the Lifecycle of Phonological Change (Bermúdez-Otero, 2007) and much of Labovian Sociolinguistics. While there is, perhaps, less emphasis on the error mechanisms triggering phonetic change in these research programs, they still commonly assume that phonologization is a gradual process.

The first step of phonologization as described by Hyman (1976, 2008) is some “intrinsic” pho-

netic effect becoming “extrinsic.”²¹ Intrinsic phonetic effects are those which are caused by natural properties of the vocal tract, just the same as those I discussed at the beginning of this chapter, and therefore subject to the same caveats. Extrinsic phonetic effects are the product of the speaker’s competence, and therefore part of either the language specific phonetics, or phonological system. Hyman (1976, 2008) does not actually argue that phonologization takes place by the gradual exaggeration of a phonetic effect, but crucially he ties the distinction between intrinsic and extrinsic phonetics to the *size* of the effect.

When the F_0 perturbations are exaggerated to a degree which cannot be attributed solely to universal phonetics, we speak of a phonologization process. (Hyman, 1976, p. 410)

The empirical fact of phonetic change, as established in Chapters 2, 4, and 5, is that the phonetic quality of vowels differentiate between contexts gradually, meaning there must be a gradual transition from the point in time where we could consider a phonetic effect to be “intrinsic” till the point in time where the phonetic effect has gotten large enough for us to consider it “extrinsic.” This conceptualization is compatible with both approaches where the boundary between phonetics and phonology is fuzzy in reality, not just for researchers, and with approaches which make a stronger assumption about qualitatively different phonology and phonetics. Under the assumption that phonology and phonetics are qualitatively different, phonologization could be conceived of as the gradual approach towards a tipping point, whereby a secondary change reinterpreting the phonetic difference of a vowel between contexts is reinterpreted as a phonological one.

Labovian sociolinguistics as a research program has traditionally made a distinction between phonological and phonetic effects, and has typically operationalized this difference in terms of the overlap of two phonetic distributions. As Labov, Karen, and Miller (1991) say, “[t]hat linguistic categories are discretely separated into mutually exclusive nonoverlapping sets is perhaps the most fundamental concept of linguistics.” This is not quite the same as the phonetic effect size metric Hyman (1976, 2008) proposes. They would classify near-mergers, for example, differently, since the size of the phonetic difference between categories is small, but so is the degree of overlap.

²¹The second stage Hyman (1976, 2008) describes involves rule inversion, and the phonemicization of phonological patterns. On this transition, Bermúdez-Otero (2007) has developed a more articulated model.

The Labovian approach to distinguishing between phonetic and phonological effects is perhaps best illustrated by the discussion of /æ/ tensing in various North American dialects in the Atlas of North American English (Labov et al., 2006, p 173–184). In particular they contrast two patterns of /æ/-tensing: the Nasal System and the Continuous System. The Nasal System has two clearly distinct allophones of /æ/. One is low and front, close to canonical [æ]. The other is longer, higher, more peripheral, and can have an inglide: [eə]. This tense-/æ/ is restricted to appear just before nasals, and the phonetic distributions of the two allophones are non-overlapping. The Continuous System is very similar, covering about the same of phonetic variation, but there is not a clear separation of allophones into non-overlapping distributions. As Labov et al. (2006, p. 180) say, however, “[i]t is evident that a continuous system of this sort differs from the nasal system only in the degree of differentiation of the vowels before nasal consonants.”

The Nasal System could be considered a phonologized version of the Continuous System, distinguished by a larger phonetic difference, and smaller phonetic overlap with /ae/ in non-nasal contexts. If a dialect with the Continuous System were to transition to a Nasal System, it would necessarily have to do so gradually, per the results of Chapters 2, 4 and 5.

6.1.2 The challenge posed by my results.

My results pose a challenge to the common and intuitive idea that conditioned phonetic change occurs due to the accumulation of production and perception errors, and that phonologization is a gradual and gradient process. First, in Chapter 4, I found that in conditioned changes where some context did not undergo the change, that context was categorically excluded from the change at its outset. Specifically, in the case of /ow/ and /uw/ fronting, these vowels before /l/ have remained unchanged, and never showed any sign of fronting along with the other contextual variants of these vowels. That is, [owl] and [uwl] allophones appear to be categorically distinguished from other allophones at the very outset of the change. Moreover, for most of the contextual variants of the vowels investigated, they moved in parallel throughout the century, even if they had very large effects. For example, the effect of a following nasal on /aw/ fronting is fairly large. In fact, around the turn of the century, the phonetic difference between [aw] and [awN] was greater than

the phonetic difference between [ow] and [owl]. Yet, the size of this effect is not predictive of their parallelism. [aw] and [awN] move in lockstep together, even beginning to reverse their trajectories together, while [ow] and [owl] begin to diverge nearly immediately.

Out of all the phonetic variants investigated, only one fits the profile of gradual divergence, and potential phonologization: /ow/ followed by nasals. Before nasals, /ow/ begins to front at about the same rate, but stalls out earlier than /ow/ in other contexts. Figure 6.3 plots these predicted trajectories from the rate of change model, along with [owl] for comparison.

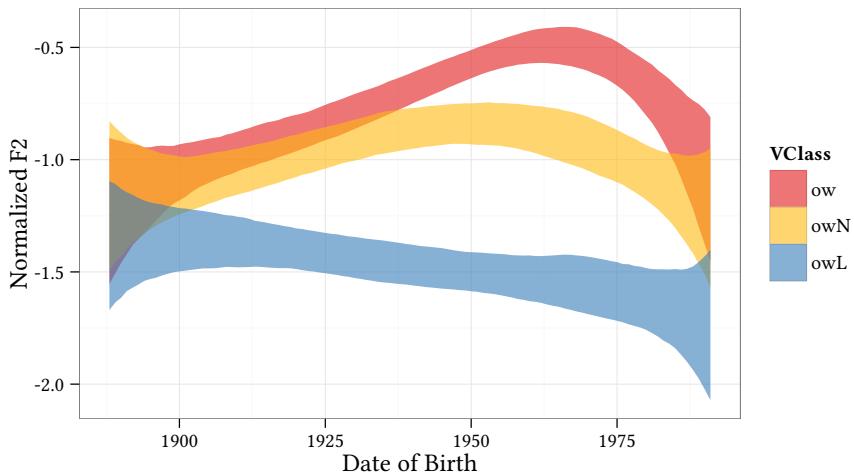


Figure 6.3: Predicted trajectories of change for /ow/ variants.

Perhaps this is the archetypal example of a phonetic process gradually becoming phonological which would be predicted under the accumulation of error model. However, it is not exactly an ideal case. Even though [owN]'s rate of change is reliably slower than [ow] (which admittedly *was* the diagnostic I proposed for distinguishing between phonetic and phonological effects), it still reaches its maximum around the same time as [ow], and even begins to retract with it. Figure 6.4 plots the predicted rates of change for [ow],[owN] and [owl], and while [owN]'s rate of change curve is in a much more compressed space than [ow], it is qualitatively very similar, especially when compared to [owl]. It seems clear that the link between [ow] and [owN] was not completely severed, as it was between [ow] and [owl], and that they are destined for similar outcomes.

If a difference between [ow] and [owN] was not phonologized, the question remains as to

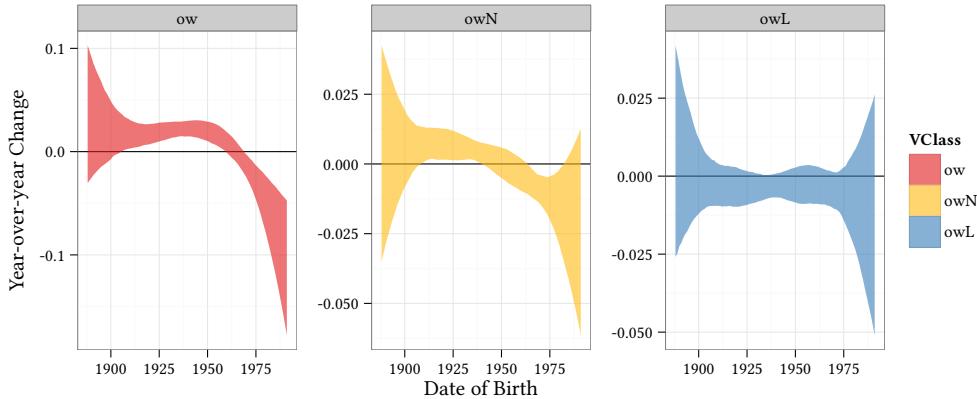


Figure 6.4: Year-to-year differences for variants of /ow/

why [owN] stalled in its fronting. The answer may be that the phonetic effect differentiating [ow] and [owN] is different from the kinds considered in Chapter 4. It should be noted that the effect a following nasal has on the F2 of /ow/ (backing) is the *opposite* of the effect it has on the F2 of /aw/ (fronting). Figure 6.5 plots density distributions in unnormalized Hz for /aw/ and /ow/, contrasting oral and nasalized variants. It appears as if the effect of nasalization biases F2 away from the vowel system center, rather than consistently in a particular direction along F2. For /aw/, which is fronting and raising, the direction away from center is essentially unbounded, allowing [aw] and [awN] to move in parallel without any apparent ceiling effects. For /ow/, on the other hand, as it fronts, it is minimizing its distance from center, perhaps amplifying the phonetic effect of nasalization. That is, the fact that [owN] slows down and stalls sooner than [ow] may also be due to a ceiling, or barrier, effect introduced by nasalization.

If we reconsider the divergence of [owN] as being due to a phonetic barrier, rather than due to a phonological reanalysis, then in fact *none* of the phonetic effects investigated in Chapter 4 became phonologized. Categorical allophones were excluded from the sound change from its very outset, i.e. they were phonologized from the very beginning. Phonetic variants moved in parallel with each other until their trajectories were perturbed by other phonetic factors, like ceiling or barrier effects.

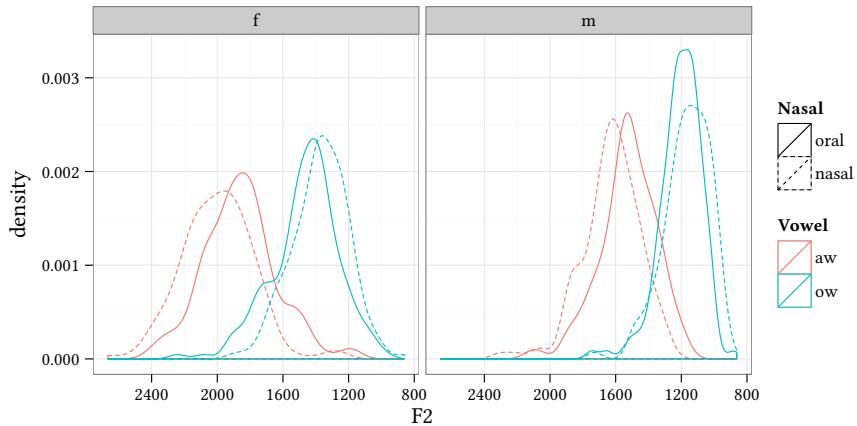


Figure 6.5: Density distributions of /aw/ and /ow/ along unnormalized F2. Dashed lines indicate nasalized vowels, solid lines oral vowels.

The results in Chapter 4 argue most strongly against a gradual process of phonologization. In Chapter 5, I found that the factors which categorize contexts as undergoing or not undergoing a change are best defined on phonological, not phonetic, grounds, at least for /ey/ and /ay/ raising. Perhaps the most surprising result is that /ay/ raising has applied opaquely with respect to flapping from the very outset of its phonetic change. Despite the demonstrable phonetic differences between surface /t/ and /d/, and their flapped forms, /ay/ raising has always applied according to the underlying voicing of the following segment. An alternative explanation to this phonological one based on lexical analogy would have to somehow take into account that /ey/ raising interacts transparently with its context. While every lexical item for /ay/ has only one or the other allophone in all contexts ([aɪ]: *ride, rider*; [ʌɪ]: *write, writer*), this is not true for lexical items for /ey/, which may have one or the other allophone depending on their context ([ɛɪ]: *pay, paying, pay off*; [eɪ]: *pays, paid, pay me*). The opaque interaction of /ay/ raising with flapping is surprising beyond just the fact that it is unexpected in the model of gradual phonologization. The Lifecycle of Phonological Change, for example, predicts that new phonological processes ought to interact transparently, and at the phrase level, like /ey/ raising does. I'll briefly discuss an analysis below which harmonizes this intuition from the Lifecycle with the /ay/ facts in Philadelphia.

The surprising result from the analysis of /ey/ raising is that the phonetic context which appeared to favor the direction of the change the most, a following /l/, did not undergo the change itself. If /ey/ in the other pre-consonantal contexts never reached the degree of fronting and raising as [eyl], it might have been possible to describe [eyl]’s non-participation in the change as a ceiling effect. However, around 1925 pre-consonantal /ey/ clearly crosses over [eyl] and continues raising, as Figure 6.6 shows. This cross over is unexpected under the accumulation of error model of phonologization. If [eyl] was higher and fronter than /ey/ in other contexts, then we should expect errors to accumulate in this context sooner and faster than in the other contexts. The non-participation of [eyl] can only be accounted for if some other factor besides its phonetic properties distinguish it from other pre-consonantal /ey/, and I argue that those properties are its phonological representation.

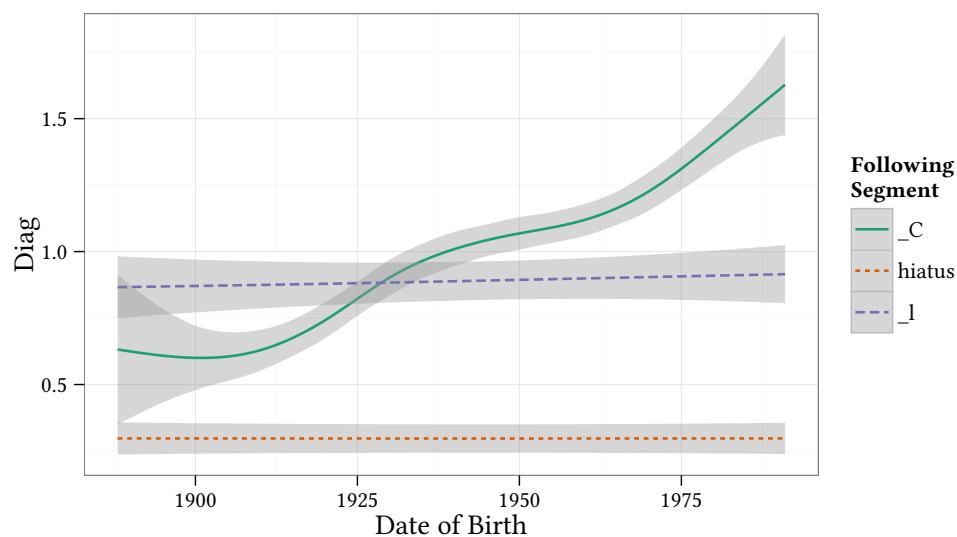


Figure 6.6: Trajectory of word internal /ey/

The preponderance of results in this dissertation so far are at least unexpected under the model of gradual phonologization. At least to the degree those models of gradual phonologization make predictions about how the process of phonologization ought to appear in diachronic data, my results have not conformed to those predictions.

Pre-Existing Phonological Processes

One possible explanation for my results which is still consistent with gradual phonologization is that in all of the cases of abrupt phonologization in this dissertation, there were actually pre-existing phonological processes in the grammar which created categorical allophones along some other phonetic dimension. For example, a different phonological process could have created two allophones of /ay/ which differentiated them along duration, and all that I observed was the further differentiation of these two allophones along an additional phonetic dimension. In fact, this is what Bermúdez-Otero (2004, p.c.) proposes to account for the surprising opacity of /ay/ raising at the outset of the change, which is either not predicted, or predicted not to be possible by the Lifecycle of Phonological Change (Bermúdez-Otero, 2007). This would mean is that I have not observed any instances of phonologization in this dissertation at all, just shifts in the phonetic realizations of pre-existing allophones.

Briefly, the Lifecycle would predict /ay/ raising to progress in the following stages, if it existed in a phonological vacuum. First, once phonologized, the new phonological process raising /ay/ to [Ai] ought to interact transparently with the surface phonology. This would predict raising in *write*, but not *writer*, and only if the /t/ wasn't flapped in a phrasal sequence like *right on*, as well as raising of word final /ay/ triggered by a following voiceless word onset. Opaque interaction with flapping, producing raising in *write* and *writer*, would come about through subsequent domain narrowing, but the exclusion of raising in phrasal context, like *lie to* would be harder to account for.

The results in Chapter 5 demonstrate fairly conclusively that this predicted sequence of historical events is *not* what happened in Philadelphia. Instead, raising always applied opaquely with respect to flapping, and the onsets of following words were never triggers for raising word final /ay/. Bermúdez-Otero's proposal is that there was a pre-existing phonological process which created two allophones of /ay/ that had the same distribution as the raising process: pre-fortis clipping (i.e. pre-voiceless vowel shortening). Pre-fortis clipping is a long standing phonological process, is present in most dialects, and crucially, as Bermúdez-Otero (2004) argues, shares the same distribution as /ay/-raising, including the opaque interaction with flapping. The argument

is that before /ay/ began to raise phonetically, there were already two phonological allophones: [ai] and [ăi], and the phonetic change raising pre-voicless /ay/ targeted only the clipped allophone, [ăi]. Whether or not an additional phonological innovation needs to be posited is an open question. Perhaps the only phonological process in the grammar is stem-level pre-fortis clipping (6.1), and all that is changing is its phonetic realization.

(6.1) ai → ăi /____ –voice]_{stem}

Or, as has been suggested to me by Bermúdez-Otero, a new phonological process is added at the phrase level which targets just [ăi].

(6.2) ăi → ăi]_{phrase}

This analysis preserves my core argument that phonetic changes operate over surface phonological representations, but does weaken the argument that phonologization is an abrupt process, because phonologization has not actually been observed in this case.

It might be possible to make a similar kind of argument for /l/ blocking the fronting of /ow/ and /uw/, because as I pointed out in Chapter 4, glide deletion for /aw/ before /l/ is a long attested feature of the Philadelphia dialect (Tucker, 1944), and the effect of /l/ on /ow/ and /uw/ could be seen as an extension of that process. For the conditioning of pre-consonantal /ey/ raising, though, it would be more difficult to propose a pre-existing phonological process. Unlike /ay/ raising, which is conditioned by just following voiceless consonants, /ey/ raising is conditioned by all following consonants, including /w/, /y/ and /r/, but not /l/. There isn't any precedent for /ey/ to split along these phonological lines reported for other dialects in the Atlas of North American English, and there isn't any other kind of phonological process I am aware of which differentiates allophones based on whether they are followed by a consonant or a glide, versus a vowel or /l/. In many ways, /ey/ raising conforms much more closely to the expectations of the Lifecycle of Phonological Change, especially in that it applies at the phrase level. Yet, it still appears to exhibit abrupt phonologization. Despite having the phonetic effect of shifting /ey/ in the direction of the change, a following /l/ never actually conditions the change itself.

It is worth considering, though, what it would mean if /ey/ raising were actually parasitic on

a previously existing phonological process. There is no evidence for such a process, but let's say that it is principled to say that one must have existed in order to explain the apparently abrupt phonologization, and that this original process entered the grammar through a mechanism of gradual phonologization. Furthermore, let's say that this is a principled explanation for any case where phonologization appears to be abrupt, which is, in fact, every case analyzed in this dissertation. The consequence would be that I have failed to observe any true instances of phonologization in this dissertation. If this is true, it would be disappointing, but would also cast doubt on the observability of phonologization. The answer to the question "What kind of data is necessary to observe phonologization?" would be "A corpus with a deeper time depth and broader coverage of the speech community than the PNC." As it is, the PNC is unparalleled in these respects, and a corpus with even an equivalent time depth and broader coverage of any speech community is unlikely to be developed any time soon.

6.2 Big Bang

My argument for an abrupt and early process of phonologization is in line with the proposal by Janda and Joseph (2003) for a "Big Bang" model of sound change, with some modifications. Their outline of their Big Bang model is quoted here in (6.3) (Janda and Joseph, 2003, (3)).

- (6.3) A "Big Bang" Theory of Sound-Change –
 - (a) sound-change originates in a very "small", highly localized context over a relatively short temporal span;
 - (b) purely phonetic conditions govern an innovation at this necessarily somewhat brief and limited point of origin;
 - (c) this brief "burst" of (an) innovation partially determines its future trajectory as it spreads through an individual's usage and through a speech community;
 - (d) the purely phonetic conditions of (b) are rapidly supplanted during spread – stage (c) immediately above – via speakers' imposition of phonological and sociolinguistic conditions, with the result that the future course of the process is thereby deflected;
 - (e) further reanalyses wholly or partially in terms of morphological and/or lexical conditions (= morpholexical – i.e., "grammatical" – ones) represent commonly occurring ultimate divergences from the initial unity of the

closely contextualized original innovation (regarding the later stages of at least one such development, see Janda 1998 on High German umlaut).

It's not exactly clear whether the examples Janda and Joseph (2003) describe can be accurately be described as being purely phonetic in origin. For example, in their example of Romance prothesis, whereby Latin word initial /sC/ clusters became /esC/ in Spanish and French, they reject "word initial" as being a possible phonetic context, because word boundaries are properly considered a phonological domain. However, they argue from evidence that the vowel prothesis was sensitive to the final segment of the preceding word, only applying if preceded by a consonant, that prothesis began "in the form of a syllable-structure-driven repair strategy." Both syllable structure and the notion of "repair" seem properly phonological. A much more phonetic explanation would probably involve something like perceptual reanalysis of consonant release as a vowel, like Blevins (2004, p. 156-7) suggests is the case for some innovations of epenthesis. It appears that for Janda and Joseph (2003), "phonetic" conditioning means something more or less like "phrase level phonology," while their process of phonologization is more analogous to domain narrowing for Bermúdez-Otero (2007), or to rule generalization.

However, Baker et al. (2011), in their critique of the error accumulation model, examine a case which does appear to go from narrow phonetic conditioning to phonological conditioning. They looked at inter-speaker variation of /s/ retraction in /str/ clusters. They classified their subjects into "retractor" and "non-retractor" groups, measured the centroid frequencies of these speakers' /s/ and /ʃ/, and compared this to the centroid frequency of /s/ in the /str/ contexts. Using ultrasound data which was also collected, they found a positive relationship between the speaker specific similarity between canonical /s/ and /r/ articulations and their degree of /s/ retraction in /str/ contexts for only the non-retractors. Their reasoning was that the more similar a speaker's /s/ and /r/ articulations are, the greater phonetic influence the /r/ should have on /s/ in /str/ contexts. The positive relationship between the articulatory similarity of /s/ and /r/ articulations and the degree of retraction in /str/ contexts suggests that for the non-retractors, /s/ retraction is simply the result of combining two independent phonetic properties in one context. For the retractors, however, there was no relationship between /s/ and /r/ articulatory similarity and

degree of /s/ retraction in /str/ contexts. Instead, these speakers produced very /ʃ/-like tokens in /str/ contexts uniformly, suggesting that /s/ retraction for these speakers is unconnected to the independent phonetic properties of their /s/ and /r/ pronunciations. In summary, they found that for some speakers, their degree of /s/ retraction was strictly proportional to independent articulatory properties, and that these speakers exhibited a broad range of phonetic variation, while for other speakers, their degree of /s/ retraction was unconnected to other articulatory properties, and that these speakers exhibited a narrower range of phonetic variation.

Baker et al. (2011) propose that the broad range of interspeaker phonetic variation among non-retractors provides the seeds for an eventual sound change. However, the sound change is not destined to happen, as it would be under the error accumulation model. Instead, they propose that the sound change leading to phonological /s/ retraction will only occur once there is an accidental alignment of speakers from less to more phonetic retraction along a relevant sociolinguistic dimension, such that speakers with more phonetic retraction are likely to be emulated.

While this proposal is attractive in that it successfully addresses the actuation problems of “Why now? Why here? Why not before or elsewhere?” as defined by Weinreich et al. (1968), it does not really address how /s/ retraction jumps from being purely phonetically conditioned to being phonologically conditioned, and unchained from the speaker specific articulations of /s/ and /r/. It is not hard to imagine at least two possibilities for how this comes about, though. The first possibility is that once the speakers with a high degree of phonetic /s/ retraction are accidentally socially situated such that most other speakers try to emulate them, the only way for speakers with little phonetic /s/ retraction to emulate them is to resort to phonological strategies. That is, the speakers with little /s/ retraction don’t naturally produce retracted /s/ in /str/ contexts, so the only way for them to emulate speakers who do is to substitute in a different phonological target, namely, the one they usually have for /ʃ/.

The second possibility is that the model of actuation proposed by Baker et al. (2011) actually requires two rare events to occur. First, some proportion of speakers must spontaneously reanalyze phonetic /s/ retraction as phonological, and second, these speakers must be socially situated such that the change spreads. This modification still relies on sporadic interspeaker variation as

the seed of change, but in this case it would be *phonological* variation, not phonetic.

This second possibility is most in line with my results, because the specific proposal from Baker et al. (2011) doesn't fit with the facts of phonetic change I'm investigating. If phonologization began first with broad phonetic variation, followed by social convergence on a particular phonetic target, we should expect the range of interspeaker variation to be very broad near the beginning of a sound change, and then begin to narrow. In fact, the interspeaker variation at the beginning ought to include in its range the eventual phonetic target that the speech community settles on. Looking at the raising of pre-voiceless /ay/, we can see that this is plainly not the case. Figure 6.7 plots the height of /ay0/ with quantile regression lines overlaid. The darkest central line represents the estimated median tendency of the speech community over time. The first set of slightly lighter lines above and below the median line are the estimated 25th and 75th percentiles over time. The outermost lines represent the 2.5th and 97.5th percentiles, so that the area in between then represents the 95% probability range of interspeaker variation. The range of interspeaker variation has remained fairly constant over time, and it was certainly not broader at the onset of the change. In fact, the 95% probability range of interspeaker variation from 1975 onwards does not overlap with the 95% range of interspeaker variation prior to 1905. The essentially constant range of interspeaker variation observed in this change remains a big mystery for the incrementation problem of how the entire speech community of Philadelphia can move in the same direction year over year, at essentially the same rate.

To recap, my proposal is that the process of phonologization appears to be more similar to the Big Bang proposed by Janda and Joseph (2003) and Baker et al. (2011) than it is to models of gradual phonetic error accumulation, like those discussed above. However, additional modifications to the Big Bang model seem to be called for on the basis of my results. The “brief” period of pure phonetic conditioning of sound change appears to be so brief as to be undetectable. In fact, this phonetic conditioning should probably not be considered part of the change itself. As Baker et al. (2011) illustrated, the phonetic conditioning of /s/ retraction was merely the product of the combination of two independent phonetic properties, and didn't really involve any innovation, in terms of a difference in linguistic competence between generations. Once an innovation is

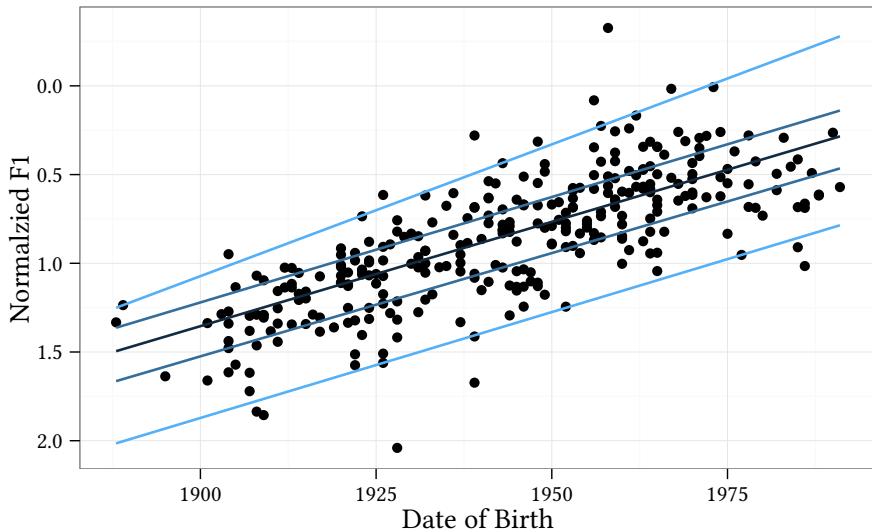


Figure 6.7: Quantile Regression over Speaker Means for [ay0]

observable, it is already phonological.

6.2.1 Plausibility

I am making two specific proposals that in this chapter, and in this dissertation, that may strain credulity.

(6.4) The initial innovation in a conditioned sound change is phonological, thus abrupt.

(6.5) The phonetic correlates of this abrupt phonological innovation are not necessarily large.

However, there is evidence in the literature on language acquisition, phonetics, phonology, and sociolinguistics which suggest that these two proposals *are* plausible.

To begin with, it may appear strange that a phonological process should appear in a speaker's grammar *ex nihilo*, out of nothing. However, if we first accept that the origins of sound changes which cannot be attributed to dialectal borrowing result largely from native language acquisition errors, then this is not so surprising. The language acquisition literature in general is dotted with examples of how children exhibit patterns divergent from the target grammar. When it comes

to phonological processes in particular, a few case studies have identified consonant harmony in children acquiring English (Smith, 1973; Pater and Werle, 2001; Gormley, 2003), a phonological process decidedly not part of the target English phonology. Of course, consonant harmony has not become a language change in progress in English, meaning that either most children abandon consonant harmony grammars before they exit the critical period of language acquisition.

Why would children adopt a phonological process for which there is no evidence in their linguistic input, only to abandon it later? Yang (2002) proposes that similar mismatches between children's syntactic grammar and the syntactic grammar generating their primary linguistic data can be attributed to their probabilistic evaluation of all possible grammars. An example from Yang (2002), most children acquiring English go through a stage of pro-drop because it is possible grammar provided by UG, and in fact most data in children's PLD is consistent with a pro-drop grammar. Only as data incompatible with pro-drop accumulates do children abandon the pro-drop grammar.

The modeling by Yang (2002) is based on a Principles and Parameters model of syntactic grammar, in which the parameters are fixed and finite. The closest existing analogy to phonology can be found in "Classical OT" (Prince and Smolensky, 2004), in which the ranking of a fixed and finite set of constraints is learned (Boersma and Hayes, 2001). However, it isn't necessary for the rules or constraints themselves to be fixed endowments of UG if instead there is some fixed and finite principles by which language learners can hypothesize new rules. Yang (2002) implicitly assumes this second possibility when modeling the acquisition of the English past tense. No one would seriously propose that a rule like $i \rightarrow æ / __T_{past}$ (*sing* → *sang*) is a primitive parameter of UG, but if we assume there are UG principles which constrain hypothesizable rules (Bergelson and Idsardi, 2009), then there is no problem in treating the probabilistic evaluation of language specific, idiosyncratic rules in a way similar to the probabilistic evaluation of UG parameters. Recently, Blaho (2008) and Samuels (2009) have made specific proposals for a set of minimal principles by which more complex and idiosyncratic phonological constraints and rules can be formulated. Blaho (2008) explicitly formulates her proposal in a model of phonological representation which is radically substance free, meaning she does not assume a universal feature set or characteristic

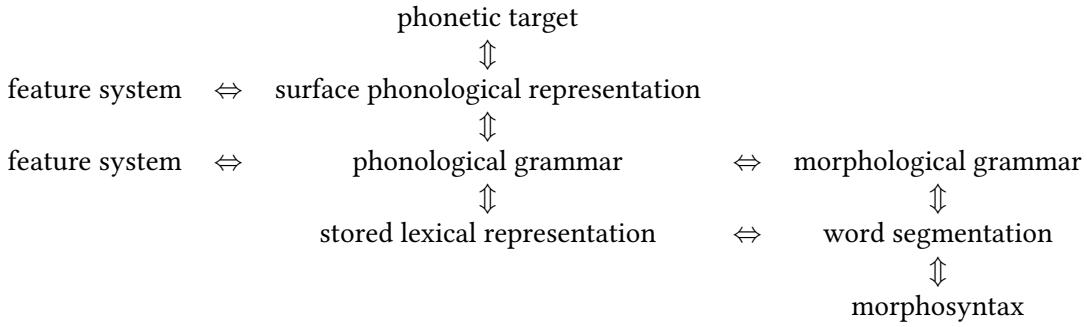


Figure 6.8: Interdependent Language Acquisition Tasks

phonetics for those features, a position to which I am sympathetic in this dissertation.

Some may still balk at the complexity of the complexity of the language acquisition task I am assuming. Not only do children need to learn the association between phonetic targets in a continuous phonetic space and categorical phonological representations, but also the phonological feature set and the set of phonological processes.²² Moreover, I'm arguing that phonological knowledge is not simply the codification of reliable phonetic patterns, so the probability of a phonological process being present in the grammar is not related to the size of observable phonetic differences. However, phonetic and phonological acquisition may be aided by the fact that they are embedded in a network of larger language acquisition tasks, sketched out in Figure 6.8. For example, Lignos (2012) proposes a model of subtractive word segmentation which crucially relies on gradually accumulating lexical representations stored in the lexicon. Of course, what the stored lexical representations *are* depends on the phonological grammar which processes the surface phonological representation. A spontaneously hypothesized process in the phonological grammar could then have the coincidental effect of boosting performance on word segmentation, which could then reinforce that process. In the same way, any new hypothesis at any location in the grammar can have a cascading reaction through this network of interdependent acquisition tasks. So while any single acquisition task may be highly complex, its interdependence on other simultaneous acquisition tasks has the effect of further narrowing the range of possibilities.

²²This is, in fact, the exact objection of (Hale and Reiss, 2008, pp. 116-7) to language specific phonetics.

In summary, my proposal is exactly that language acquirers can hypothesize new phonological processes *ex nihilo* because they can freely generate hypotheses which then compete. This has the interesting result that Bermúdez-Otero's (2007) observation that most new phonological processes apply at the phrase level is generalization without as strong an explanation. One possibility, though, may be that most hypothesized phonological processes which become language changes are hypothesized early in phonological acquisition, before learners have mastered word segmentation.

Errors in native language acquisition are likely to be distributed sporadically throughout a speech community, however, and are likely to be highly idiosyncratic compared to something like syntactic acquisition. Looking at the literature, *some* children acquiring English are reported to hypothesize consonant harmony, while most children undergo a stage where they pro-drop some proportion of the time (Yang, 2002). The question arises, then, how a phonological process that a language learner spontaneously hypothesizes becomes a language change, and more importantly, how the entire speech community could suddenly possess the same phonological process. In answer to this, I think it is instructive to look at a case where a speech community has *not* converged on the same phonological process.

Mielke et al. (forthcoming) examine idiosyncratic differences between speakers' articulation of /r/: bunched or retroflex. The articulatory difference between bunched and retroflex /r/ is large, but there is no, or minimal, difference in their acoustics. While most of their subjects were either categorical bunched or retroflex /r/ users (16 out of 27), the remaining subjects exhibited variation between the two variants. Mielke et al. (forthcoming) observe a considerable amount of idiosyncratic constraints on the distribution of bunched versus retroflex /r/. In total, they propose 22 constraints in order to account for the distribution of /r/ variants across all speakers in their study, and of those 22 constraints, 13 (13/22 = 59%) were represented by only a single speaker, and 86% (19 out of 22) were represented by three or fewer speakers. However, some constraints were more common. For example, the constraint *RETROFLEX/CODA, which disallows retroflex /r/ in codas, was the most common, present in 9 out of 11 variable speakers.

The fascinating fact about Mielke et al. (forthcoming) is that speakers who vary between bunched and retroflex /r/ have any structured constraints regarding their distributions at all. Because there are not reliable acoustic cues to which articulation is being used, there is no data available to learners as to which variant they are hearing in any given case. But rather than sporadically distribute bunched and retroflex /r/ across all contexts, speakers appear to develop internally consistent grammars. In some sense, the distribution of /r/ articulations is a constrained version of “the forbidden experiment” of language deprivation. Language learners are deprived of information about which option to take for a decision which is not strongly constrained by UG principles, and the result is internal consistency, but with a high rate of idiosyncratic variation when speakers are compared. Even more fascinating is that *some* patterns are more common than others, a fact that Mielke et al. (forthcoming) attributed to general tendencies in American English for exaggerated anterior gestures in syllable onsets, like that observed for light versus dark /l/ (Sproat and Fujimura, 1993).

We know from language acquisition research that children will spontaneously hypothesize phonological processes, and from Mielke et al. (forthcoming) that these phonological processes may not result in any measurable phonetic difference, *and* that in effective isolation, speakers will spontaneously hypothesize the *same* phonological process. As Ringe and Eska (2013) point out

In any major city in the world there must be at least tens of thousands of children in the [native language acquisition] developmental window at any given time. If only one child in a thousand persists in a learner error until the period of [native language acquisition] is past, that type of event will be too rare to be recognizable in any sociolinguistic survey, yet there will be a steady stream of new variants brought into the speech community as the children grow up.

And importantly, drawing from the results presented by Mielke et al. (forthcoming), many language learners will spontaneously hypothesize and persist in the *same* “error,” or mismatch from the grammars of the previous generation. In order to persist into a speaker’s adult grammar, following the logic of Yang (2002), all the newly hypothesized phonological process needs to do is not lose, meaning there just has to be little enough data inconsistent with it.

In order to be identified as a language change, the new phonological process must diffuse throughout the speech community, and that process imposes its own narrowing effects on the

change. Following the reasoning of Ringe and Eska (2013) and Baker et al. (2011) and the results of Mielke et al. (forthcoming), I'll suggest that there is a constant stream of children with idiosyncratic phonological grammars surrounding some potential innovation, and that occasionally the distribution of speakers with the innovation will coincidentally correlate with sociolinguistic dimensions which promote its spread through the speech community. Labov (2010b) argues that language learning is largely outwardly oriented, meaning children are socially motivated to coordinate their grammars to conform to their peer group. Citing work by Payne (1980) and Kerswill and Williams (2000), he argues that children abandon the models of their parents in favor of their peer group. This tendency to conform to the consensus of the peer group would, in most cases, eliminate the idiosyncratic phonological innovation of any single individual, which is why it is necessary to propose that in order for a language change to take place, it would have to be independently innovated by many children, which again appears to be plausible given the results of Mielke et al. (forthcoming).

This outward orientation of language acquisition may also play a role in the small phonetic correlates of categorical phonological innovation. The fact that Mielke et al. (forthcoming) found phonological variation which correlated with nearly uniform acoustics is, I believe, enough of a plausibility test to demonstrate that phonological differences don't necessarily correspond to large acoustic differences. But with the case of bunched versus retroflex /r/, the fact that there is no acoustic difference is due to the fact that the two articulations produce the same acoustics, and we might not expect this to be the case for all phonological innovations. Moreover, a shift from bunched to retroflex /r/, or vice versa, has not become a change in progress, for the very reason that there is no acoustic difference for speakers to attend to.

However, let's say that a learner hypothesizes a phonological process which creates two allophones of /ay/: [ay₁] and [ay₂]. The speaker now has to decide what the phonetic realizations of [ay₁] and [ay₂] ought to be. If they're living in a speech community for which most speakers have only one allophone of /ay/, then the best way to conform their two allophone grammar with the broader speech community is to decide that [ay₁] and [ay₂] have very similar phonetics targets. This is similar to the argument that Dinkin (2011b) makes for the backing of short-o (the LOT

vowel) in Upstate New York. Most of Upstate New York participates in the Northern Cities Shift, which includes the fronting of short-o towards [a] or [æ]. However, Upstate New York is bordered to the South by Western PA, to the North and West by Canada, and to the East by Northern New England, all of which have the low-back merger of /ɑ/ and /ɔ/. Dinkin (2011b) finds that in Upstate New York, the phonetic difference between /ɑ/ and /ɔ/ has been decreasing in response, he hypothesizes, to contact with merged dialects. This mirrors the famous “Bill Peters Effect,” Labov (1994) whereby a speaker living in a merged dialect region still produced a reliable phonetic difference for the phonemic contrast between /ɑ/ and /ɔ/ in free conversation, but produced them merged in a minimal pairs task. My conclusion on this point is that even if there were a natural tendency for language learners to posit large phonetic difference to go along with phonological differences, these phonetic differences could get reduced by sociolinguistic homogenization.

6.2.2 Big Bang Summary

In conclusion, my results are more in line with a “Big Bang” model of conditioned sound change in which phonological innovations occur at the *onset* of the change, rather than as a reanalysis later on. Both the facts that this means that speakers are innovating a new phonological process *ex nihilo*, and that towards the beginning of this change the phonological innovation corresponds to a small phonetic difference are plausible given what we know about language acquisition, phonology, phonetics and sociolinguistics.

6.3 Similarity to syntactic change.

It is worth noting that a debate between gradual versus abrupt phonologization closely mirrors a similar discussion in syntactic change. Hyman (2008, p 398-9) actually draws the connection between “phonologization” and “grammaticalization,” drawing the four part analogy “phonetics : phonology :: pragmatics : syntax.” In a review of grammaticalization and gradualness, Traugott and Trousdale (2010) describe the basic position on gradualness in much of the grammaticalization literature:

Gradualness refers to the fact that most change involves (a series of) micro-changes, an issue which is sometimes overlooked in considerations of more general patterns of language change. As Brinton and Traugott (2005: 150) observe, although change is sometimes understood (or at least formulated) as A > B, studies of gradualness in linguistic change attempt to uncover “the tiny local steps between A and B that the arrow ‘>’ encompasses”.

This is very similar to Kroch’s summary of the field of historical syntax in 1989.

The idea that language change proceeds context by context, with new forms appearing first in a narrowly restricted context and spreading to others only later, has been widely accepted. It has seemed obvious that the ordering of contexts in the spread of a change reflected the linguistic forces causing the change.

Of course, Kroch (1989) was arguing against this position on the basis of the evidence of the constant rate effect. Instead, he argued, syntactic change is abrupt and catastrophic, meaning all possible contexts are included in the scope of the change at its onset. More recently, Denis (2013) has made the same argument for the distribution of a change across pragmatic contexts. Denis (2013) examined the frequency of use of utterance final particles (UFP) (e.g. *right, you know*), and found that even though younger speakers appear to use the new UFP, *right*, in a broader range of pragmatic contexts than older speakers, this appearance is strictly modulated by their baseline usage frequency of *right*. That is, in his data, the fact that older speakers are only observed to use *right* in 2 out of 10 possible pragmatic contexts is quantitatively indistinguishable from the hypothesis that they can and do use *right* in all possible pragmatic contexts, but they use *right* at such a low frequency to begin with that it would take more data than is feasible to collect to observe them doing so.

The results in my dissertation further cement the position of Fruehwald et al. (forthcoming) that the mechanisms of phonological and syntactic innovation are fundamentally similar. Innovation in both sound change and syntactic change is abrupt, and does not take place through the gradual reanalysis of phonetic or pragmatic phenomena, respectively. In both cases, after the original innovation, most of the observed change involves either increasing phonetic differentiation or increasing frequency of use of the innovation. There is no reason why sound change and syntactic change *must* have been subject to the same dynamics, but it does appear that they are.

6.4 Additional Challenges, and Directions for Future Research

There are a number of interesting research questions which I have been unable to address in this dissertation which I will have to reserve for future research. For example, I believe that the difference between phonological allophones and phonetic variants ought to have broader sociolinguistic consequences than I have been able address here. I conclude that the difference between [ow] and [owl] is phonological in origin, while the difference between [aw] and [awN] is phonetic. From this, I would assume that it is possible for [ow] and [owl] to have disconnected stylistic usage, while it would be impossible for [aw] and [awN]. That is, a speaker could not raise [awN] to an extreme level for a stylistic purpose that they could not also raise [aw], whereas a speaker could front [owl] for a stylistic purpose which is separate from the stylistic fronting and backing of [ow]. It has already been established that pre-voiceless /ay/ has this property, where the backing of [ay0] indexes masculinity and toughness while the frontness or backness of low [ay] has not been reported to have any similar indexical purpose (Conn, 2005; Wagner, 2007). At the moment, however, this reasoning is completely speculative, and requires more careful studies of stylistic variation which take the distinction between phonological and phonetic variants into account.

Another interesting direction of research would be to investigate how phonological allophony for one vowel can influence others. Labov (2010a), for example, argues that allophonic chain shifting is impossible. To support this argument, he looks at dialects which have a large difference between pre-nasal and pre-oral /æ/. In those dialects where pre-nasal /æ/ is extremely raised and fronted, he finds no concomitant fronting of pre-nasal /a/. However, in the Northern Cities Shift, where /æ/ is uniformly raised and fronted, there is concomitant fronting of /a/ to [a] or [æ]. Labov (2010a) attributes the lack of allophonic chain shifts to the “Binding Force” of segmental phonology, or the dictum that “phonemes change.” While the absence of allophonic chain shifting cast doubt on loosely structured phonemic representation like those proposed by Exemplar Theory, they may not be entirely impossible. Many North American dialects exhibit a phonological process distinguishing pre-nasal and pre-oral /æ/, but none have been reported to do something similar for /a/, so while there may be two phonological allophones of /æ/ ([ə̃] and [æ]), there is

only one for /ɑ/. The fact that pre-nasal /ɑ/ does not front in reaction to pre-nasal /æ/ raising could simply be because there is no relevant allophone to front. If Labov's assertion that allophonic chain shifting is impossible is true, a more complex model of the phonology-phonetics interface than the one I've pursued here will be necessary. Specifically, the interface will need to define a relationship between a phonetic target, a surface phonological representation *and* its phonemic identity. On the other hand, if an example of an allophonic chain shift is discovered, then the explanation for the lack of pre-nasal /ɑ/ fronting in the dialects Labov (2010a) investigated will have to shift to the distinction between phonological and phonetic variation that I just described.

A more important problem to address, however, is the challenge of identifying phonological innovation, specifically the assertion of Ringe and Eska (2013) that it will be “too rare to be recognizable in any sociolinguistic survey.” However, if it is the case, as I have argued, that the phonological innovations which become sound changes are those innovations which multiple speakers are likely to independently produce, then they shouldn't be *impossible* to detect. The focus and methodology of the sociolinguistic surveys aimed at detecting these innovations may need to be adjusted. For example, the Peaks model of language change incrementation (Labov, 2001; Tagliamonte and D'Arcy, 2009) places most of the action in incrementation squarely on adolescents, specifically between the time they first enter their peer groups and the end of adolescence. Focusing on this demographic of speakers will be necessary to identify new phonological innovations.

Chapter 7

Conclusions

In this dissertation, I have set out to understand how phonology and phonetics interact over the course of phonetic change with the hope of broadening our knowledge of both sound change, and the general relationship between phonology and phonetics. Drawing upon the data in the Philadelphia Neighborhood Corpus, I've been able to examine the time course of phonetic change in fine detail, and arrived at some surprising results.

- (7.1) When a context is observed to exert a categorical effect (either categorically conditioning or blocking) a phonetic change, that categorical effect is typically already in place at the onset of the change (Chapter 4).
- (7.2) Robust phonetic effects were rarely, or never reanalyzed as being phonological (Chapter 4).
- (7.3) The way in which contexts behave as triggers or non-triggers of phonetic changes are frequently best described in phonological, rather than phonetic terms (Chapter 5).
 - (a) The raising of pre-voiceless /ay/ has always patterned according to the underlying voicing of the following segment, not its surface phonetic realization.
 - (b) The raising of pre-consonantal /ey/ was never conditioned by following /l/, even though a following /l/ appeared to always phonetically favor that change.
- (7.4) There is a striking parallelism across vowel categories for many changes, even after taking into account their social correlation (Chapter 5).

On the basis of these results, I arrive at two primary conclusions. First is that the phonetic changes I have observed in this dissertation operate over surface phonological representations. In terms of the grammatical model I outlined in Chapter 2, phonetic changes are the shifting phonetic implementation of surface phonological representations. It may be possible for alternative grammatical models to explain the results that I found, but they do not *predict* them. Secondly, I argue that phonological innovations are not the *product* of phonetic change, but rather are in place at onset of phonetic changes. This conclusion is a rather substantial shift away from the conventional wisdom regarding phonologizaton, which is why I devote a considerable portion of Chapter 6 to arguing for the plausibility of this conclusion.

There are a number of ways in which this research project can be pushed forward. First and foremost, in this dissertation I have examined 7 vowel shifts in Philadelphia. In order for my results to be maximally credible, replications of these results for as many changes in as many speech communities as possible will be necessary. Moreover, these results should hold true of other kinds of phonetic change beyond just shifts in a vowel's central tendency.

Secondly, there are a number of points I made in the argument for the plausibility of spontaneous phonologization which require further investigation. Specifically, a broader search for idiosyncratic phonological variation which corresponds to small phonetic correlates, like those discovered by Mielke et al. (forthcoming), should be carried out. A key demographic group to turn to for such a search would be adolescents. According to the Peaks Model (Labov, 2001; Tagliamonte and D'Arcy, 2009), adolescents undergo active reorganization of their language from those forms acquired on the basis their caregiver input, to those further influenced and constrained by the social structure of their peer groups. It would be ideal to capture the idiosyncratic phonological variation, if it exists, at this period of speakers' life, before they are ironed out by social forces.

Ideally, studies of this kind could also be added to the toolkit of phonological investigation. Just as the fact that a set of segments pattern together either as triggers or undergoers of a phonological process can be taken as evidence for their membership in a phonological natural class, so can the fact that they all undergo the same phonological change. In addition, principles like the

Unity Principle, which I proposed in Chapter 4, can be used to differentiate between phonological and phonetic processes.

Bibliography

- Adank, Patti, Roel Smits, and Roeland van Hout. 2004. A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America* 116:3099. URL <http://link.aip.org/link/JASMAN/v116/i5/p3099/s1&Agg=doi>.
- Anttila, Arto, Vivienne Fong, Štefan Beňuš, and Jennifer Nycz. 2008. Variation and opacity in Singapore English consonant clusters. *Phonology* 25:181. URL http://www.journals.cambridge.org/abstract_S0952675708001462.
- Ash, Sherry. 1982. The Vocalization of /l/ in Philadelphia. Ph.d., University of Pennsylvania.
- Bailey, Charles-James. 1973. Variation and Linguistic Theory. Technical report, Center for Applied Linguistics, Washington.
- 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. URL http://www.journals.cambridge.org/abstract_S0954394511000135.
- Baranowski, Maciej. 2008. The fronting of the back upgliding vowels in Charleston, South Carolina. *Language Variation and Change* 20:527–551. URL http://www.journals.cambridge.org/abstract_S0954394508000136.
- Bates, Douglas. 2006. [R] lmer, p-values and all that. URL <https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html>.
- Becker, Kara. 2010. Regional Dialect Features on the Lower East Side of New York City: Socio-phonetics , Ethnicity , and Identity. Ph.d., New York University.
- Becker, Kara, and Amy Wing-mei Wong. 2010. The Short-a System of New York City English: An Update. *University of Pennsylvania Working Papers in Linguistics* 15. URL <http://repository.upenn.edu/pwpl/vol15/iss2/3/>.
- Bergelson, Elika, and WJ William Idsardi. 2009. Structural Biases in Phonology: Infant and Adult Evidence from Artificial Language Learning. In *BUCLD 33: Proceedings of the 33rd annual Boston University Conference on Language Development*, 85–96.
- Bermúdez-Otero, Ricardo. 2004. Raising and Flapping in Canadian English: grammar and acquisition. In *Handout of paper presented at the CASTL Colloquium, University of Tromsø. 2, November 2004*. URL <http://www.bermudez-otero.com/tromsoe.pdf>.

- Bermúdez-Otero, Ricardo. 2007. Diachronic Phonology. In *The Cambridge Handbook of Phonology*, ed. Paul de Lacy, chapter 21, 497–517. Cambridge: Cambridge University Press.
- Blaho, Sylvia. 2008. The Syntax of Phonology: A radically substance free approach. Ph.d., University of Tromsø.
- Blevins, Juliette. 2004. *Evolutionary phonology: the emergence of sound patterns*. New York: Cambridge University Press.
- Blevins, Juliette. 2006. A theoretical synopsis of Evolutionary Phonology. *Theoretical Linguistics* 32:117–166. URL <http://www.degruyter.com/view/j/thli.2006.32.issue-2/tl.2006.009/tl.2006.009.xml>.
- Blevins, Juliette, and Andrew Garrett. 2004. The Evolution of Metathesis. In *Phonetically Based Phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade, chapter 5, 117–156. New York: Cambridge University Press.
- Boberg, Charles. 2005. The Canadian shift in Montreal. *Language Variation and Change* 17:133–154.
- de Boer, Bart. 2000. Self-organization in vowel systems. *Journal of Phonetics* 27:441–465. URL <http://www.sciencedirect.com/science/article/pii/S0095447000901256>.
- de Boer, Bart. 2001. *The Origins of Vowel Systems*. New York: Oxford University Press.
- Boersma, Paul, and Silke Hamann. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25:217–270.
- Boersma, Paul, and Bruce Hayes. 2001. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32:45–86.
- 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. URL <http://repository.upenn.edu/pwpl/vol17/iss1/5/>.
- Browman, Catherine, and Louis M Goldstein. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3:219–252.
- Buckley, Eugene. 2011. Metathesis. In *The Blackwell Companion to Phonology, Volume 3*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, volume 1968. Blackwell.
- Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14:261–290.
- Bybee, Joan. 2007. Word Frequency in Lexical Diffusion and the Source of Morphophonological Change. In *Frequency of use and the organization of language*, ed. Joan Bybee, chapter 2, 23—34. New York: Oxford University Press.
- Bybee, Joan, and James L McClelland. 2005. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review* 22:381–410.

- Chambers, J K. 1973. Canadian Raising. *Canadian Journal of Linguistics* 18:131–135.
- Chládková, Kateřina, and Silke Hamann. 2011. High Vowels in Southern British English: /u/-Fronting does not result in merger. *Proc. 17th ICPHS* 17–21. URL <http://www.icphs2011.hk/resources/OnlineProceedings/RegularSession/Chladkova/Chladkova.pdf>.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- 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 C. 1993. Nasalisation in English: Phonology or Phonetics. *Phonology* 10:43–81.
- Cohn, Abigail C. 2006. Is there Gradient Phonology? In *Gradience in grammar: Generative perspectives*, ed. Gisbert Fanselow, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky, 25–44. Oxford: Oxford University Press.
- Cohn, Abigail C. 2007. Phonetics in Phonology and Phonology in Phonetics. *Working Papers of the Cornell Phonetics Laboratory* 16:1–31.
- Conn, Jeffery. 2005. Of ‘moice’ and men: The evolution of a male-led sound change. Ph.d., University of Pennsylvania.
- Coye, Dale F. 2010. Dialect Boundaries in New Jersey. *American Speech* 84:414–452. URL <http://americanSpeech.dukejournals.org/cgi/doi/10.1215/00031283-2009-032>.
- Dailey-O’Cain, Jennifer. 1997. Canadian raising in a midwestern U.S. city. *Language Variation and Change* 9:107–120.
- Darlington, Richard B. 1970. Is Kurtosis Really “Peakedness?” *The American Statistician* 24:19–22.
- Denis, Derek. 2013. Grammaticalization? Change in the right periphery from 1879 to 2003 Development of UFPs In Ontario English. Paper Presented at the 87th Annual Meeting of the Linguistic Society of America.
- Dinkin, Aaron. 2011a. What’s Really Happening to Short A Before L in Philadelphia? Paper presented at NWAV40.
- Dinkin, Aaron J. 2011b. Weakening resistance: Progress toward the low back merger in New York State. *Language Variation and Change* 23:315–345. URL http://www.journals.cambridge.org/abstract_S0954394511000147.
- Disner, Sandra Ferrari. 1978. Vowels in Germanic Languages. Ph.d., UCLA.
- Docherty, Gerard J, and Paul Foulkes. 2000. Speaker, Speech, and Knowledge of Sounds. In *Phonological knowledge: Conceptual and empirical issues*, ed. Noel Burton-Roberts, Philip Carr, and Gerard J Docherty. Oxford: Oxford University Press.
- Dresher, B. Elan. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.

- Dresher, B. Elan, Christopher Harvey, and Will Oxford. 2012. Contrast Shift as a Type of Diachronic Change.
- Durian, David. 2012. A New Perspective on Vowel Variation across the 19th and 20th Centuries in Columbus, OH. Ph.d., The Ohio State University.
- Ellegård, Alvar. 1953. *The auxiliary do: The establishment and regulation of its use in English*. Stockholm: Almqvist and Wiksell.
- Espy-Wilson, Carol, and Suzanne Boyce. 1994. Acoustic differences between "bunched" and "retroflex" variants of American English /r/. *The Journal of the Acoustical Society of America* 95:2823.
- Ettlinger, Marc. 2007. An Exemplar-based model of chain shifts. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, ed. Lawrence Erlbaum, 239–244.
- Evanini, Keelan. 2009. The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania. Ph.d., University of Pennsylvania.
- Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18:7–44.
- Flemming, Edward. 2004. Contrast and preceptual distinctiveness. In *Phonetically Based Phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade, 232–276. Cambridge University Press.
- Foulkes, Paul, James M Scobbie, and Dominic Watt. 2010. Sociophonetics. In *The handbook of phonetic sciences*, ed. William J Hardcastle, John Laver, and Fiona E Gibbon, 703–754. Blackwell.
- Fourakis, Marios, and Robert Port. 1986. Stop Epenthesis in English. *Journal of Phonetics* 14:197–221.
- Fridland, Valerie. 2001. The social dimension of the Southern Vowel Shift: Gender, age and class. *Journal of Sociolinguistics* 5:233–253. URL <http://doi.wiley.com/10.1111/1467-9481.00149>.
- Fruehwald, Josef. 2008. The Spread of Raising : Opacity , Lexicalization , and Diffusion. In *Penn Working Papers in Linguistics*, ed. Kyle Gorman, volume 14, 83–92. URL <http://repository.upenn.edu/pwp1/vol14/iss2/11/>.
- Fruehwald, Josef, Jonathan Gress Wright, and Joel Wallenberg. forthcoming. Phonological Rule Change: The Constant Rate Effect. In *The proceedings of the North-Eastern Linguistic Society (NELS)*. URL http://www.ling.upenn.edu/~joseff/papers/FGW_CRE_NELS40.pdf.
- Gafos, Adamantios I, and Stefan Benus. 2006. Dynamics of phonological cognition. *Cognitive Science* 30:905–943. URL http://onlinelibrary.wiley.com/doi/10.1207/s15516709cog0000_80/abstract.
- Garrett, Andrew, and Keith Johnson. 2011. Phonetic bias in sound change. *UC Berkely Phonology Lab Annual Report* 9–61.

- Gorman, Kyle, Jonathan Howell, and Michael Wagner. 2011. Prosodylab-Aligner: A tool for forced alignment of laboratory speech. In *Proceedings of Acoustics Week in Canada, Quebec City*, 4–5.
- Gormley, Andrea Lucienne. 2003. The Production of Consonant Harmony in Child Speech. Ma, Concordia University.
- Gress-Wright, Jonathan. 2010. Opacity and Transparency in Phonological Change. Ph.d., University of Pennsylvania.
- Guy, Gregory R. 2007. Lexical Exceptions in Variable Phonology. *University of Pennsylvania Working Papers in Linguistics* 13:109–119.
- Haddican, Bill, Paul Foulkes, Vincent Hughes, and Hazel Richards. forthcoming. Social correlates of two vowel changes in Northern England Social correlates of two vowel changes in Northern England. *Language Variation and Change*.
- Hale, Mark. 2004. Neogrammarian Sound Change. In *The Handbook of Historical Linguistics*, ed. Brian D Joseph and Richard D Janda. Blackwell.
- Hale, Mark, Madelyn Kissock, and Charles Reiss. 2007. Microvariation, variation, and features of universal grammar. *Lingua* 117:645–665.
- Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. New York: Oxford University Press.
- Hall, Daniel Currie. 2007. The Role and Representation of Contrast in Phonological Theory. Ph.d., University of Toronto.
- Hayes, Bruce, and Donca Steriade. 2004. Introduction: the phonetic bases of phonological Markedness. In *Phonetically based phonology*, ed. Bruce Hayes, Robert Kirchner, and Donca Steriade. Cambridge University Press.
- Henderson, Anita Louise. 2001. Is your money where your mouth is? Hiring managers' attitudes toward African-American Vernacular English. Ph.d., University of Pennsylvania.
- Herold, Ruth. 1990. Mechanisms of merger: The implementation and distribution of the low back merger in eastern Pennsylvania. Ph.d., University of Pennsylvania.
- Hoenigswald, Henry. 1960. *Language change and linguistic reconstruction*. University of Chicago Press.
- Hoffman, Matthew D, and Andrew Gelman. 2011. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *arXiv.org* URL <http://arxiv.org/abs/1111.4246>.
- Hume, Elizabeth. 2000. Metathesis Website. URL <http://www.ling.ohio-state.edu/~ehume/metathesis/>.
- Hyman, Larry M. 1976. Phonologization. In *Linguistic studies offered to Joseph Greenberg on the Occasion of His Sixtieth Birthday*, ed. Alphonse Juillard, A.M. Devine, and Laurence D. Stephens, volume 4, 407–418. Anima Libri.

- Hyman, Larry M. 2008. Enlarging the scope of phonologization. *University of California Berkeley Phonology Lab Annual Report* 382–409. URL http://www.linguistics.berkeley.edu/PhonLab/annual_report/documents/2008/Hyman_Phonologization_PLAR.pdf.
- 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 Papers from the 15th International Conference on Historical Linguistics, Melbourne*, ed. Barry Blake and Kate Burridge, 205–219. John Benjamins.
- Johnson, Daniel Ezra. 2007. Stability and change along a dialect boundary: The low vowels of Southeastern New England. Ph.d., University of Pennsylvania.
- Joos, Martin. 1942. A phonological dilemma in Canadian English. *Language* 18:141–144.
- Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. Ph.d., MIT.
- Keating, Patricia A. 1985. Universal phonetics and the organisation of grammars. In *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, ed. Victoria A Fromkin, 115–132. New York: Academic Press.
- Keating, Patricia A. 1988. Underspecification in phonetics. *Phonology* 275–292.
- Keating, Patricia A. 1990. Phonetic representations in a generative grammar. *Journal of phonetics* .
- Kerswill, Paul, and Ann Williams. 2000. Creating a New Town koine : Children and language change in Milton Keynes. *Language in Society* 29:65–115. URL <http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=28334>.
- Kiparsky, Paul. 1965. Phonological Change. Ph.d., Massachusetts Institute of Technology.
- Kiparsky, Paul. 2006. The Amphichronic Program vs. Evolutionary Phonology. *Theoretical Linguistics* 32:217–236. URL <http://www.degruyter.com/view/j/thli.2006.32.issue-2/tl.2006.015/tl.2006.015.xml>.
- Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.
- Kroch, Anthony. 1994. Morphosyntactic Variation. In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. K Beals.
- Kruschke, John K. 2011. *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. Oxford: Elsevier.
- Kwong, Katherine, and Kenneth Stevens. 1999. On the voiced-voiceless distinction for writer/rider. *Speech Communication Group Working Papers (Research Laboratory of Electronics at MIT)* 11:1–20.

- 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*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change. Volume 2: Social Factors. Language in Society*. Oxford: Blackwell.
- Labov, William. 2006a. Plotnik 08. Retrieved September 2006 <http://www.ling.upenn.edu/~wlabov/Plotnik.html>.
- Labov, William. 2006b. *The social stratification of English in New York City*. New York: Cambridge University Press, second edition.
- Labov, William. 2007. Transmission and Diffusion. *Langauge* 83:344–387.
- Labov, William. 2010a. *Principles of Linguistic Change. Volume 3: Cognitive and Cultural Factors*. Oxford: Blackwell.
- Labov, William. 2010b. What is to be learned? Paper presented at NELS 41.
- Labov, William, Sherry Ash, and Charles Boberg. 2006. *The Atlas of North American English*. New York: Mouton de Gruyter.
- Labov, William, and Maciej Baranowski. 2006. 50 Msec. *Language Variation and Change* 18:223–240.
- Labov, William, David Graff, and Wendell A. Harris. 1986. Testing listeners' reactions to phonological markers. In *Diversity and Diachrony*, ed. David Sankoff, 45–58. Philadelphia: John Benjamins.
- Labov, William, Mark Karen, and Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3:33–74. URL http://journals.cambridge.org/abstract_S0954394500000442.
- Labov, William, and Ingrid Rosenfelder. 2011. The Philadelphia Neighborhood Corpus.
- Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear Incrementaion, Reversal, and Reanalysis. *Language* 89:30–65.
- Labov, William, Malcah Yaeger, and Richard Steiner. 1972. *A Quantitative Study of Sound Change in Progress. The U.S. National Survey*.
- Liberman, Mark, and Janet B Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In *Language sound structure: studies in phonology presented to morris halle*, ed. Mark Aronof and R T Oehrle, 157–233. Cambridge: MIT Press.
- Lignos, Constantine. 2012. Infant Word Segmentation: An Incremental, Integrated Model. In *Proceedings of the 30th West Coast Conference on Formal Linguistics*, ed. Natan Arnett and Ryan Bennett, 237–247. Somerville, MA: Cascadilla Proceedings Project.

- Liljencrants, Johan, and Björn Lindblom. 1972. Numerical Simulation of Vowel Quality Systems : The Role of Perceptual Contrast. *Language* 48:839–862.
- MacKenzie, Laurel, and Meredith Tamminga. 2012. Non-local Conditioning of Variation: Evidence and Implications.
- Martinet, André. 1952. Function, structure and sound change. *Word* 8:1–32.
- Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. New York: Oxford University Press.
- Mielke, Jeff, Adam Baker, and Diana Archangeli. forthcoming. A typology of idiosyncratic /ɪ/ allophony patterns in American English.
- Mohanan, K. P. 1992. Describing the phonology of non-native varieties of a language. *World Englishes* 11:111–128. URL <http://doi.wiley.com/10.1111/j.1467-971X.1992.tb00056.x>.
- Moreton, Elliott, and Erik R Thomas. 2007. 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, 37–64. Cambridge University Press.
- Odden, David. 2006. Phonology ex nihilo. Paper presented at the the Tromsø Phonology Project Group Meeting.
- Ohala, John J. 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.
- Ohala, John J. 1990. The phonetics and phonology of aspects of assimilation. In *Papers in laboratory phonology 1: Between the grammar and physics of speech*, ed. John Kingston and Mary E Beckman, 258–282. Cambridge: Cambridge University Press.
- Oxford, Will. 2012. Patterns of contrast in phonological change : Evidence from Algonquian vowel systems. URL http://individual.utoronto.ca/woxford/Oxford_Ms_PatternsOfContrast.pdf.
- Parrell, Benjamin. 2012. The role of gestural phasing in Western Andalusian Spanish aspiration. *Journal of phonetics* 40:37–45. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3381366/>&tool=pmcentrez&rendertype=abstract.
- Pater, Joe, and Adam Werle. 2001. Typology and variation in child consonant harmony. In *Proceedings of HILP*, ed. Caroline Féry, Antony Dubach Green, and Ruben van de Vijver. University of Potsdam. URL <http://people.umass.edu/pater/pater-werle1.pdf>.
- Payne, Arvilla. 1980. Factors Controlling the Acquisition of the Philadelphia Dialect by Out-of-State Children. In *Locating Language in Time and Space*, ed. William Labov. New York: Academic Press.
- Phillips, Betty S. 1984. Word Frequency and the Actuation of Sound Change Betty S. Phillips. *Language* 60:320–342.

- Phillips, Betty S. 1999. The mental lexicon: evidence from lexical diffusion. *Brain and language* 68:104–9. URL <http://www.ncbi.nlm.nih.gov/pubmed/10433746>.
- Phillips, Betty S. 2006. *Word Frequency and Lexical Diffusion*. Basingstoke: Palgrave Macmillian.
- Pierrehumbert, Janet B. 1990. Phonological and phonetic representation. *Journal of Phonetics* 18:375–394.
- Pierrehumbert, Janet B. 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.
- Pierrehumbert, Janet B. 2002. Word-specific phonetics. In *Laboratory Phonology VII*, 101–139. Mouton de Gruyter.
- Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34:516–530. URL <http://linkinghub.elsevier.com/retrieve/pii/S009544700600043X>.
- Preston, Dennis R. 2004. Three kinds of sociolinguistics: A psycholinguistic perspective. In *Sociolinguistic variation: Critical reflections*, ed. Carmen Fought, 140—158. New York: Oxford University Press.
- Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.
- R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL <http://www.r-project.org/>.
- Ringe, Don, and Joseph F. Eska. 2013. *Historical Linguistics: Toward a Twenty-First Century Reintegration*. New York: Cambridge University Press.
- Rosenfelder, Ingrid. 2005. Canadian Raising in Victoria , B. C.: an Acoustic Analysis. Ma, Ruprecht Karl University of Heidelberg.
- Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, and Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite. URL <http://fave.ling.upenn.edu/>.
- Ruch, Hanna. 2012. Investigating a gradual metathesis: Production and perception of /s/ aspiration in Andalusian Spanish.
- Samuels, Bridget D. 2009. The Structure of Phonological Theory. Ph.d., Harvard University.
- Sankoff, Gillian. 2006. Gillian Sankoff. Age: Apparent time and real time. In *Encyclopedia of language and linguistics*. Oxford: Elsevier.
- Sankoff, Gillian, and Hélène Blondeau. 2007. Language Change across the Lifespan: /r/ in Montreal French. *Language* 83:560–588.
- Scobbie, James M. 2005. The phonetics phonology overlap. *QMUL Speech Science Research Centre Working Papers*, WP-1 1–30.

- Simpson, Adrian P. 2009. Phonetic differences between male and female speech. *Language and Linguistics Compass* 3:621–640. URL <http://doi.wiley.com/10.1111/j.1749-818X.2009.00125.x>.
- Smith, Neilson V. 1973. *The Acquisition of Phonology: A Case Study*. New York: Cambridge University Press.
- Sproat, R, and O Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of phonetics* 21:291–311.
- Stan Development Team. 2012. Stan: A C++ Library for Probability and Sampling. URL <http://mc-stan.org/>.
- Stevens, Kenneth. 1989. On the *quantal* nature of speech. *Journal of Phonetics* 17:3–45.
- Stevens, Kenneth N, and HM Hanson. 2010. Articulatory-Acoustic Relations as the Basis of Distinctive Contrasts. In *The Handbook of Phonetic Sciences*, ed. William J. Hardcastle, John Laver, and Fiona E. Gibbon, 424–453. Blackwell, second edition.
- Strycharczuk, Patrycja. 2012. Phonetics-phonology interactions in pre-sonorant voicing. Ph.d., University of Manchester.
- Tagliamonte, Sali, and Alexandra D'Arcy. 2009. Peaks Beyond Phonology: Adolescence, Incrementation, and Language Change. *Language* 85:58–108.
- Tamminga, Meredith. 2012. Persistence in the production of linguistic variation.
- Torreira, Francisco. 2006. Coarticulation between Aspirated-s and Voiceless Stops in Spanish : An Interdialectal Comparison. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 1978, 113–120. Somerville, MA: Cascadilla Proceedings Project.
- Torriera, Francisco. 2007. Pre- and postaspirated stops in Andalusian Spanish. In *Segmental and prosodic issues in Romance phonology*, ed. P. Prieto, J. Mascaró, and M.-J. Solé, 67–82. Philadelphia: John Benjamins.
- Traugott, Elizabeth Closs, and Graeme Trousdale. 2010. Gradience, gradualness and grammaticalization: How do they intersect? In *Gradience, gradualness and grammaticalization*, ed. Elizabeth Closs Traugott and Graeme Trousdale, volume i, 19–44. Amsterdam: John Benjamins.
- Tucker, R Whitney. 1944. Notes on the Philadelphia dialect. *American Speech* 19:37–42.
- Wagner, Suzanne E. 2007. We act like girls and we don't act like men: The use of the male-associated variable (ay0) in South Philadelphia. In *Penn Working Papers in Linguistics*, ed. Tatjana Scheffler, Joshua Tauberer, Aviad Eilam, and Laia Mayol, volume 13.1, 393–406. Philadelphia: University of Pennsylvania Press.
- Warner, Anthony. 2005. Why DO dove: Evidence for register variation in Early Modern English negatives. *Language Variation and Change* 17:257–280.

- Watt, Dominic J. L. 2000. Phonetic parallels between the close-mid vowels of Tyneside English: Are they internally or externally motivated? *Language Variation and Change* 12:69–101. URL http://www.journals.cambridge.org/abstract_S0954394500121040.
- 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. 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*.
- Zimman, Lal. 2013. Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language and Sexuality* 2:1–39. URL <http://openurl.ingenta.com/content/xref?genre=article&issn=2211-3770&volume=2&issue=1&spage=1>.
- Zsiga, Elizabeth C. 2000. Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28:69–102.