Second Dialect Acquisition: Implications for Theories of Phonological Representation

by

Jennifer R. Nycz

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

New York University

January, 2011

Gregory R.Guy

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge the Canadians in New York and New Jersey who generously gave their time and their vowels to this project; without them, there would be no data, and no dissertation! Thanks to all of my speakers for sharing their stories, which were at turns gripping, reflective, funny, insightful, heart-rending, and often some combination of the above.

The next big thank-you goes to **Greg Guy**. I was very fortunate to find an advisor whose intellectual passions, love of puns, and fierce prioritization of meal-times coincided so exactly with my own. I've had countless chats with Greg about THE BIG ISSUES IN LINGUISTICS over the years, and in addition to informing and inspiring the work presented here, nearly all of those discussions were just plain *fun*. All errors in this dissertation are mine, as the platitude goes; but if, dear reader, you find yourself raising an eyebrow at some of the more playful turns of phrase contained herein, you can at least partially blame Greg, because he only encouraged that sort of thing.

I am grateful to **John Singler**, whose knowledge, guidance, banter, and merciless wielding of the editing pen (never in red) has immeasurably improved the quality of both my work and my overall graduate school experience. His final-quarter coaching is what finally got me to the endzone.

Thanks also to the rest of my committee - **Diamandis Gafos, Lisa Davidson** and **Gerry Docherty** - for their thought-provoking and helpful feedback on this work.

These five professors are the only ones who have actually had to read my dissertation, but many others have encouraged and supported me throughout its creation and my linguistic career.

I've probably learned the most from day-to-day interactions with my fellow grad students at NYU. Thank you to **Kevin Roon, Cara Shousterman, Marcos Rohena Madrazo, Tuuli Morrill, Jason Shaw, Libby Coggshall, Stacy Dickerman, Kara Becker, Amanda Dye, Danny Erker, Tricia Irwin, Maryam Bakht, Lisa Levinson, Jon Brennan, Oana Savescu Ciucivara, Mike Taylor, Ken Lacy, Tom Leu, Sonya Fix, Simanique Moody, Laura Rimell** and anyone else I overlapped with for your willingness to share your stimulating conversation and general hilarity, often over plates of something delicious. Thanks also to **Renee Blake, John Costello, Ray Dougherty, Chris Barker, Anna Szabolcsi, Mark Baltin,** and all of the other faculty at NYU for contributing to a lively, collegial academic environment.

At Dartmouth College, **Lindsay Whaley** and **Ioana Chitoran** were the first to show me how exciting Linguistics could be, and graciously tolerated all the jokes (in my defense: Principles & Parameters? Comedy *gold*). Later, ☞ **Arto Anttila** provided a solid foundation in phonology & variation, while **Doug Honorof** at Haskins Laboratories introduced me to much of the experimental literature discussed in these pages. **Dominic Watt**, **Carmen Llamas**, **Sam Hellmuth** and

Paul Foulkes at the University of York have also given useful comments, advice, and copious amounts of coffee.

The ambient Canadian vowels of **Paul De Decker** are what originally sparked this line of research; phonemic inventory aside, Paul has been a good friend and a most excellent (and patient!) co-author.

Daniel Ezra Johnson, the original Dissertation Buddy$^{\text{TM}}$, was especially helpful during the early days of this project, and has offered challenging discussion and statistical advice throughout the thesis-writing (and non-writing) years.

Maria Gouskova started out as my boss, but quickly turned into a friend and valued mentor. One of my proudest realizations about finishing the doctorate is that we can now exchange terrible YouTube videos *as colleagues*.

Special thanks go to **Damien Hall**, who made the difficult double-shift (full-time RAship + PhD writing) bearable with his cheery office demeanor and sixth sense for when I needed a cuppa. Thanks also to **Lauren Hall-Lew** for the virtual (and, lucky me, occasional IRL!) empathy and company on this side of the pond, to **Laura Gibson** for the cheerleading and the Ferragosto cds which always seemed to arrive at just the right time, to **Dave Nycz** for the absorbing and distracting (in a good way) brotherly gchats about politics, math, and cats, and to **Steve Nycz** for his continued interest in his niece's work as well as the general encouragement he's given since I was just a curious little kid.

Holden Chi Hoon Lee, my oldest close friend, has been a comforting constant in a decade which often has seemed far too full of variables.

Amy Wong, NWAV Roommate For Life, buoyed me with her special mix of positivity, love, and bluntness, and kept me up-to-date on all the North American linguistics news.

Lisa Roberts, a ray of sunshine in northern England, offered cheery chats, soothing chocolate, and pictures of fluffy things when especially needed.

Eytan Zweig and I were chums at NYU, but it was in York that we became quite close. Words cannot adequately express how grateful I am to have such a supportive, understanding friend, so instead of attempting to thank him in prose, I will just give him a cake.

Stuart Brown has been my study buddy, sparring partner, pub pal, kvetching crony, hug-giving homie, and all-around main mate. He's the best thing about England (yes, even counting Twiglets!), and he helps in so many ways. I really could not have finished this work without him.

Finally, thank you to my parents **Kathy and Frank Nycz**, who may not have understood what I've been doing all these years, nor really believed in the existence of people who think *cot* and *caught* sound the same, but have always supported my choices.

ABSTRACT

How do speakers who move to a new dialect region acquire features of the new dialect? Social factors surely affect this process; for example, the degree to which a speaker wishes to align with the new community will modulate how features associated with that community are acquired. However, linguistic factors - the form of phonological representations, their malleability, and the processes that manipulate them to yield surface forms - must also constrain the types of variation and change available to the individual speaker. This dissertation sets out the predictions made by generative phonology and usage-based phonology regarding how such change should occur, and uses second dialect acquisition data to test these predictions.

The study draws its data from sociolinguistic interviews with mobile adults who acquired their native dialect of English in Canada and later moved to the New York City region. It focuses on the linguistic and social factors affecting acquisition of two phonological variables which differ across these two regions: the *cot/caught* distinction and height of (aw) in Canadian Raising environments. A sociophonetic analysis of these variables was undertaken to determine whether each speaker has acquired New York-like realizations of these vowels, and whether this acquisition seemed to be occurring on a lexically gradual basis. The relationship between these features across speakers was also examined.

Several findings emerge from this study. Most of the speakers in the sample have acquired a *cot/caught* distinction after years spent in the New York region, but maintain a raised (aw) nucleus, especially in salient lexical items such as *about*; however, both features show evidence of phonetically and lexically gradual shift as predicted by usage- based theory. A positive correlation was found between degree of *cot/caught* distinction and degree of Canadian Raising: those speakers with the greatest distance between *cot* and *caught* words are also those who exhibit the most raised (aw) diphthongs. I argue that these findings support a model in which phonological representations are both phonetically rich and linked to social labels, and propose the addition of a new parameter to the model which accounts for the correlation between the two features.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.0 Overview

Human beings are compulsive imitators, prone to unconsciously mimicking the postures and gestures of those with whom they interact (LaFrance 1982, Bernieri and Rosenthal 1991, Cappella 1981, 1996, Chartrand and Bargh 1999, Shockley et al. 2003). This is especially true in the realm of speech, where each person's innate ability to imitate allows him to acquire a first language as a child and ultimately fine-tune his speech to match the dialect(s) spoken around him (Condon and Sander 1974, Kuhl and Meltzoff 1996). While this ability seems to lessen with age, most mature speakers continue to demonstrate some degree of accommodation in conversation, partially and temporarily shifting both global features of voice quality (Natale 1975, Gregory 1990) and pronunciation of particular words and sounds (Bell 1984, Coupland 1984, Giles 1973, Giles et al. 1991, Pardo 2006) to become more similar to those of an interlocutor. However, speakers are also capable of more permanent shifts: some individuals, over time, seem to change aspects of their accent when embedded in a new speech community. Which linguistic features are subject to accommodation and which not? When do ephemeral shifts give way to lasting change in an individual's language? And finally, why do some changes occur and not others?

Social factors are surely an important part of the answer: the degree to which a speaker wishes to align herself with a given interlocutor or with a more abstract identity will modulate the extent to which features associated with that interlocutor or identity are incorporated into the speaker's language. However, **linguistic factors** - the form of phonological representations, their malleability, and the processes that manipulate them to yield surface forms - must also constrain the types of variation and change available to the individual speaker. Data on intraspeaker change therefore has the potential to shed light on foundational questions in phonological theory.

This dissertation uses just this sort of data to address just these sorts of questions, through a sociolinguistic study of mobile adults: specifically, native speakers of Canadian English who have moved to the New York/New Jersey region, and have thus been exposed to new dialect forms. I will examine how these speakers realize two phonological features which differ across these dialects, to determine whether (and to what extent) speakers have acquired new dialect variants for each. These data will be analyzed with the goal of testing specific predictions made by **generative phonology** and **usage-based phonology**, two prominent models which make very different claims about phonological representations and processes.

This chapter begins by situating the dissertation within the broader context of the interface between variation and phonological theory, before moving on to a discussion of the empirical and theoretical impetus for the study and a review of the relevant background literature.

1.1 Variation & phonological theory

In his seminal work on New York City English, Labov endorses a sociolinguistics which uses "data from the speech community to solve problems of linguistic theory" (Labov 1966, viii). In the nearly fifty years since that foundational publication, sociolinguists working within the Labovian framework have engaged with and contributed to phonological theory to varying extents.[1] The most obvious contribution of this research is a large and continually growing body of empirical evidence that there are systematic linguistic constraints on variation (Weinreich et al. 1968, Labov 1994), indicating the need for a phonological theory that can account for this variation.

However, sociolinguists and sociolinguistic data have also played a more direct role in the development of such a theory. The earliest variationist foray into formal phonology was Labov (1969)'s introduction of the *variable rule*, a modification of the context-sensitive rewrite rule described by Chomsky and Halle (1968) in The Sound Pattern of English (henceforth, SPE). In SPE, rules which variably apply are marked [+optional], with application or non-application of the rule ultimately "depending on the outcome of the algorithm responsible for optionality" (Vaux 2008, p. 42). This formalization is essentially a handwave, yet it makes the nontrivial prediction that linguistic context (nevermind social context!) has no effect on rule application: because the optionality algorithm presumably operates

[1] Sociolinguists have also grappled with theoretical issues in morphosyntax (see e.g. Kroch 1994, Montgomery 1994, Tagliamonte 1998). However, this dissertation will focus on phonological variation.

at a level 'above' rules, it can only 'see' the feature [OPTIONAL], so the structural description that partially composes each rule can have no effect on its outcome. This prediction, of course, is false: linguistic context does affect the probability of rule application. The variable rule framework accommodates this fact by locating the optionality mechanism in the rule itself: each contextual factor which conditions rule outcome is associated with a particular probability weight, and all of these weights contribute to the overall probability of rule application. While this innovation was never embraced by mainstream formal phonology (Coetzee and Pater, To Appear), the variable rule and its implications were further developed by sociolinguists such as Cedergren and Sankoff (1974) and Sankoff and Labov (1979), who focused on the underlying statistical model, and Guy (1991a,b), who argued that variable rules could operate cyclically within Lexical Phonology (Kiparsky 1982).

This interaction between variationist sociolinguistics and phonology continued with the rise of Optimality Theory (OT) (Prince and Smolensky 1993). OT was initially developed to account for the variation *between* languages, via the variable ranking of a universal set of constraints. However, scholars soon realized its potential for dealing with both the qualitative and quantitative facts of variation *within* languages. In fact, the OT era arguably ushered in the first period of true bilateral talks between sociolinguists and phonologists. Linguists working in the Labovian paradigm continued to incorporate the developments of phonology into their research, adapting OT to account for the phenomena they had always been interested in (e.g. Reynolds 1994, Nagy and Reynolds 1997, Zubritskaya 1997, Auger 2001, Cardoso 2001, Nycz 2006) and testing the new theory against older models (Myers 1994, Guy 1997), but with the introduction of OT, formal phonologists also began to look to variable data to test and refine their theories. Anttila (1997, 2002a), for example, uses variation in the Finnish genitive plural to argue for an OT model in which constraints may be partially ordered by the grammar: for any particular evaluation of a candidate set, a specific total order consistent with this partial ordering is randomly chosen; variation results when different output candidates are judged optimal by the various total orderings. Coetzee (2004, 2006), in contrast, draws on vowel lenition in Faialense Portuguese and other variable phenomena to support his theory that variation results instead from the harmonic rank-ordering imposed on the candidate set, with less harmonic candidates being available as dispreferred variants. Boersma and Hayes (2001), meanwhile, propose a Stochastic OT in which constraints are ordered along a continuous ranking scale, and show how such a model is capable of modeling fine-grained quantitative variation, using metathesis and reduplication in Ilokano as test cases.

Variationist data has thus been important to the development of theories of phonological *processes* since the 1970s. At one level, this has involved adapting models (whether rule- or constraint-based) to account for both the qualitative and quantitative patterns found in specific data sets. However, deeper questions have also been raised by this interaction between variation and phonological theory. For example, what is the relationship between categorical and variable phenomena: are they fundamentally distinct conceptual objects, treated differently by the grammar and subject to different learning strategies during acquisition (Dell 1981, Vaux 2008), or do they exhibit similar constraints and patterns, indicating the need for an integrated account (Guy and Boberg 1997, Boersma and Hayes 2001, Anttila 2002b, Bakovic and Pajak 2008)? How are the effects of extragrammatical factors such as frequency (Coetzee 2009) and style (van Oostendorp 1997) best accommodated in the grammar, if at all? Finally, in a framework (such as OT) in which inter- and intra-language variation receive the same theoretical treatment, is there a principled distinction to be made between the two?

In contrast, there has been very little engagement between sociolinguistics and phonology with respect to the development of theories of phonological *representation*.[2] For decades, sociolinguistic work has largely assumed a theory of representation similar to that presented in SPE. In this framework, the phonological portion of a lexical entry consists of a highly abstract sequence of segments composed of distinctive features, with the phonetic details of how these items are actually produced filled in at later stages of implementation. Of course, variationists and phonologists have always differed in terms of the relative theoretical importance they ascribe to the various components of the SPE model: the details of phonetic implementation, while trivia to Chomsky and Halle, are important indicators of social affiliation and linguistic change to sociolinguists. However, up through the 1990s, the model itself had never been subject to much reflection or problematizing within the pages of Language Variation and Change.

More recently, scholars working within the variation-friendly field of Laboratory Phonology (Pierrehumbert et al. 2000) have argued that the information

[2]It is telling that review articles dealing with 'Variation and Phonological Theory' are almost entirely discussions of phonological processes. Anttila (2002b)'s chapter in the Handbook of Language Variation and Change covers general issues raised by variation and presents several OT models of variable phenomena, but contains no mention of phonological representation. Similarly, Coetzee and Pater's contribution to the Handbook of Phonological Theory (2nd ed.) addresses the question of where to locate variation in phonology and provides an overview of both OT and variable rule approaches, but is nearly silent on representational issues; while the authors briefly acknowledge the existence of Exemplar Theory as an alternative to standard generative assumptions about representation, they do not engage with this alternative. Coetzee and Pater do address the problem of lexically conditioned variation, noting that frequency effects and lexical exceptions suggest the need for lexically indexed constraints in the grammar; however, this is not a claim about phonological representation, but about how processes may be related to (the same old) representations.

stored within lexical entries is much richer than previously thought, drawing on theories of memory and representation developed in psychology. The best-known version of this idea in linguistics is Exemplar Theory (ET), which holds that "each category is represented in memory by a large cloud of remembered tokens of that category" (Pierrehumbert et al. 2000, 3); these categories can include allophones, phonemes, words, frequent collocations, and even social groupings. New tokens of these categories are derived from their associated exemplar clouds.

Many argue that ET is an especially promising model for dealing with sociolinguistic data, since it handles gradient phonetic variation while also linking this variation to social categories (Mendoza-Denton et al. 2003, Hay et al. 2006a,b, Hay and Drager 2007, Foulkes and Docherty 2006, Drager 2009, Foulkes 2010). However, mainstream sociolinguists, while acknowledging that ET might offer advantages at a certain level of explanation, have more or less maintained their generative assumptions (Labov 2006b, e.g.). This is unfortunate, because sociolinguistic data can provide excellent test cases for deciding between these theories of representation. In the next section, I discuss a case in point: linguistic accounts of Herzog's Principle.

1.2 Herzog's Principle: Competing accounts

One robust finding in the field of dialectology is that phonological mergers tend to spread at the expense of distinctions (Herzog 1965, Labov 1994). Labov (1994) christened this observation "Herzog's Principle," based on Herzog's study of mergers affecting high vowels in the Yiddish spoken in Northern Poland. Several changes which have been observed in North American varieties of English are consistent with this principle, including the northern and westward spread of the *pin/pen* merger beyond its Southern dialect origins (Brown 1990, Labov et al. 2006), the decline of the w/ʍ contrast in pairs like *which/witch*, and the merger of or/ɔr in words like *four/for* (Labov 1994). One especially well-studied instantiation of this principle is the low back vowel merger, "the largest single phonological change occurring in American English" (Labov 1994, 316). While the Inland North, the South, the Eastern Midlands, and the New York City dialect regions continue to contrast the vowels in the word classes exemplified by *cot* /kɑt/ and *caught* /kɔt/,[3] these words are no longer distinct in Eastern New England, Western Pennsylvania, the West, and Canada (Labov et al. 2006). Moreover, current research has shown that the merger is appearing in several other areas, including Missouri (Gordon 2001, Majors 2005), Kentucky (Irons 2005), and West Virginia (Hazen 2003).

Why is it that one type of phonological pattern spreads easily, while its converse does not? This issue is typically framed as a question about contact between *dialects*, where a dialect is an abstract linguistic system shared by many speakers and thus separate from any speaker; these reified language varieties are discussed as if they float like warm and cold air masses over their associated regions, with contact-induced linguistic changes occasionally raining down at the fronts. However, the actual entities coming into contact with one another are not *dialects* of American English, but individual *speakers* who embody similar-yet-not-identical linguistic systems, and change through accommodation is possible whenever two such speakers interact. A complete account of Herzog's Principle, or any other observation pertaining to the diffusion of dialect features, will therefore require an

[3]For dialects of American English which contrast these vowels, these words are typically phonemically transcribed with /ɑ/ and /ɔ/, respectively, though the precise phonetic realization of each of these vowels differs across dialect regions.

understanding of the constraints affecting individual speakers who are exposed to non-native dialects.

So, why would individuals who natively contrast two sounds abandon, rather than retain, their distinction when faced with merged interlocutors? Why would the merged speakers fail to acquire a contrast in the same contact situation? Social factors are potentially relevant: the distinction may be avoided if it (or the dialect it characterizes) receives a negative social evaluation, or if the merger is seen as more standard. However, while it makes sense to consider the effects of social evaluation in particular instances of merger, such factors cannot explain Herzog's Principle in general, because it is possible that in some cases the distinction (or its associated dialect) will receive a more positive social evaluation than the corresponding merger.[4]

Some linguists (e.g. Herold 1990, Labov 1994) have offered accounts of the asymmetry in terms of the linguistic constraints on change over a speaker's lifespan. As described in Labov (1994), the *cot/caught*-contrasting, two-phoneme (henceforth, 2P[5]) speaker who is exposed to a one-phoneme (1P) dialect need only "ignore" his native distinction in order to accommodate to this dialect. While it is unclear exactly what status the action of "ignoring" a contrast has in linguistic

[4]One possible example of this case is the spread of the *pin/pen* merger. A feature of Southern speech (which is, to put it mildly, not a positively evaluated variety of American English (Preston 1999)), this merger has nonetheless spread to parts of the South Midlands and Western dialect regions. But even in the absence of specific counterexamples like this one, there would be little reason to think that contrast per se would be socially stigmatized relative to merger; indeed, the prescriptive importance attributed to "speaking clearly and distinctly" would lead one to expect the opposite.

[5]Throughout this dissertation I will describe speakers as being natively 2P and 1P, to reflect the number of low back vowel phonemes which each type of speaker is assumed to have. I will avoid the usual descriptions of "distinct" and "merged" for individual speakers, so as not to imply any synchronic process of merger taking place in the latter group (There is no reason to believe that "merged" speakers from Canada, for instance, have distinct phonological representations for *cot* and *caught* which are ultimately neutralized; they simply have one low back vowel phoneme).

theory, one way to translate this into formal phonological terms is to say that such a speaker need only apply a single neutralization rule to his distinct underlying representations in order to generate the right output. The 1P speaker, however, has a much more complicated task if she is to learn the distinction: for every word in her lexicon which contains a [+LOW +BACK] vowel specification, she must now ascertain whether that word is additionally [+ROUND] or [-ROUND], [+TENSE] or [-TENSE], or whatever the relevant contrastive feature may be. This "internal" account crucially assumes a classical, feature-based theory of phonological representations, in which underlying forms consist of abstract symbols with little phonetic substance, which are later spelled-out by phonetic implementation rules that treat all instances of a given symbol the same way; in such a theory, new surface contrasts can only be realized if new symbols marking this contrast have been added to the relevant underlying representations.

As discussed in the previous section, however, there exist competing linguistic theories that make different predictions regarding how new forms may be incorporated into a speaker's grammar or lexicon. According to ET, if a 1P speaker - whose exemplar clouds for *cot* and *caught* largely or completely overlap in phonetic space - is exposed to a dialect in which *cot* and *caught* are pronounced with different vowel sounds, his representations of these words should gradually start to diverge, ultimately resulting in productions of these words that also contrast. Similarly, the 2P speaker's *cot* and *caught* are expected to gradually converge with exposure to a dialect in which these words are pronounced the same. If this is in fact what happens, then the observation that mergers tend to spread over dialect regions at the expense of contrasts cannot readily be explained by invoking an asymmetrical speaker-level constraint on contrast acquisition.

It is clear, then, that a post hoc explanation of Herzog's Principle which appeals to the tenets of classical phonological theory is merely one hypothesis which needs to be tested against competing hypotheses, such as that offered by the Exemplar Theoretic account. However, this fact is not often recognized in the literature on dialect contact. This amounts to theoretical and empirical lost opportunities: dialect variation and change data are indeed relevant to formal theoretical debates, and appropriately scientific attempts to resolve theoretical questions should push researchers to seek out new kinds of data.

This dissertation will pursue exactly this type of research program, by studying dialect contact and intraspeaker change with the explicit goal of testing different theories of representation. In the remainder of this chapter, I will review relevant literature on intraspeaker change and accommodation from multiple fields, focusing on findings that may be relevant to the theoretical questions at hand as well as the techniques used to gather the data, then outline the approach taken in the current study.

1.3 Previous work on intraspeaker change

The general question of how speakers change their language over the lifespan has been addressed in the sociolinguistic and the psychology/laboratory phonology literature from different theoretical and methodological perspectives. Here I review some of the most relevant literature from each field, beginning with sociolinguistic studies of change over the lifespan in speakers who remain in one speech community. I then consider studies of language change by speakers who have moved to a new speech community characterized by a different dialect. Finally, I cover laboratory studies of the fine-grained changes that speakers make after short-term exposure to speech that differs from their own.

1.3.1 Change over the lifespan

Change over the lifespan has not always been a central concern in sociolinguistics. This is largely due to what Labov (2006a) refers to as "the central dogma of sociolinguistics: that the community is prior to the individual" (pg 5); while idiolects considered in isolation may be characterized by puzzling irregularities and exceptions ("studded with oscillations and contradictions", in the words of Labov 1966), such wrinkles tend to be ironed out in a larger group analysis, where systematic patterning according to social and stylistic factors is revealed. This focus on uncovering more general patterns of socially conditioned variation, combined with the related goal of understanding how language as an "abstract pattern, exterior to the individual" (Labov et al. 2006, 5) changes over time, does not leave much room for detailed studies of the individual. Moreover, sociolinguistic work on language change in apparent time is predicated on the idea that individuals do not, in fact, change their language very much after adolescence, in keeping with Lenneberg (1967)'s critical period hypothesis. Finally, there are practical issues associated with the longitudinal study of individual speakers, such as subject attrition and limitations on time and funding, which make intraspeaker change studies more difficult to carry out (Tillery and Bailey 2003).

Despite these impediments, there are a handful of sociolinguistic studies investigating phonological change over the lifespan. Some of these projects have used real-time panel data to validate the results of concurrent or previous apparent-time studies. For example, Sankoff and Blondeau (2007) used real- and apparent-time data to examine the change in realization of /r/ in Montreal French during the latter part of the 20th century. Sankoff and Blondeau's 1971 data showed evidence for a change in progress in apparent time, with an apical [r] variant preferred by older speakers seemingly giving way to a dorsal [R] preferred by younger speakers. To examine the extent to which this feature might also be changing in the speech of individuals, Sankoff and Blondeau looked more closely at the use of (r) by 32 speakers who were recorded in both 1971 and 1984. In 1971, most of the speakers showed a strong preference for one variant or the other: 12 were majority [r] users (using the innovative form no more than 17% of the time) and 10 were majority [R] users (using this variant at least 85% of the time), with the remaining 10 speakers characterized as mid-range users. Comparing these data with those from 1984, Sankoff and Blondeau found that the majority of speakers were stable in their use of the variable, though 9 of their speakers did significantly shift their speech in the direction of using more of the innovative [R]. Moreover, most of these shifting speakers (7/9) were from the group of 1971 midrange users.

Similarly, Nahkola and Saanilahti (2004) used a panel study alongside an apparent-time study to track changes in 14 morphological and phonological variables in a rural variety of Finnish. There was much idiolectal variation among the 24 speakers in their panel, though Nahkola and Saanilahti found that the majority of variables remained stable over time. However, where a particular variable did shift for an individual, this shift tended to be in the direction of the community change. Like Sankoff and Blondeau, Nahkola and Saanilahti observe that, on an individual speaker basis, it is the variables with no categorical or near-categorically-preferred variant which seem most likely to change over time.

Other researchers have carried out longitudinal studies with the goal of examining change over the lifespan of individuals per se. A recent example is Bowie (2009)'s study of phonological changes in the speech of Mormon church leaders, as evidenced in recordings of speeches made by these men over the course of several decades. Bowie looked at two phonological variables, (ʍ)-voicing and realization of coda (r), and found much interspeaker variation: while every speaker in his sample changes his use of both of these features over time, there is no overall pattern or directionality to these changes. However, Bowie notes that neither of these variables have been recognized as features of Utah (or even Western American) English, and do not seem to be participating in either stable variation or change in progress in this area. There is therefore no external, community-based target for these speakers to be shifting in relation to.

Other studies of individuals, however, have focused on features which *have* been implicated in broader changes in progress. For example, Harrington et al. (2000) and Harrington (2006) looked at fine-grained phonetic shift in the speech of Queen Elizabeth II, by conducting acoustic analysis of vowel tokens taken from the Queen's Christmas broadcasts spanning a 50 year period. Harrington observed changes in the Queen's vowel productions over time, such that the shape of her overall vowel space and the position of the HAPPY-vowel shifted to reflect wider community developments. De Decker (2006) followed up with a group of 4 girls from his 2002 study of phonetic changes in a small Ontario town, to track their use of (æ)-retraction, a feature of the Canadian Shift. He rerecorded the girls 3 years after they had moved to Toronto,[6] where the Canadian shift is more advanced,

[6] While the speakers in De Decker's study did move from their hometown to the city, this is arguably not a case of second-dialect contact. The variety associated with Toronto is not qualitatively different from that of the small Ontario town the girls were raised in, simply more advanced with respect to changes affecting both localities.

and found that 3 of the 4 girls had increased their use of retracted (æ) - that is, had changed to reflect the advancement of the speech around them.[7] Wagner (2008) assembled a somewhat larger panel of 9 female high school students in Philadelphia, analyzing their use of 5 phonological variables in 2005 and again in 2006. Two of these variables, (dh) and (ing), are stable in the community, and likewise remained mostly stable in the speech of the girls, with only the highest socioeconomic group decreasing their use of the nonstandard variants for each variable. However, realizations of (ay0), (aw), and (e) did change over time in these individuals, reflecting ongoing vowel shifts in Philadelphia.

The conclusions emerging from this small collection of studies is that individuals who remain in more or less the same linguistic community may alter their use of phonological variables to reflect ongoing community changes, and this seems to be most likely when the variables in question are already unstable in the individual's speech. That is, speakers are capable of making quantitative changes in their phonological behavior to reflect ongoing quantitative changes in the community. What happens, however, when speakers move to a new community and are exposed to a *qualitatively* different system?

1.3.2 Acquisition of second dialect features

Before becoming a research question in its own right, the effect of second dialect exposure on a speaker's vernacular was considered an obstacle to be overcome or avoided in traditional dialectological work. These studies focused on the speech of non-mobile older rural males ("NORMs"), whose lack of contact with non-native

[7] De Decker notes that mere exposure to the community change is not enough to promote individual change. Attitude and orientation to the community is also important: the one speaker who did not increase her use of retracted (æ) otherwise stood out from the rest of the panel by not participating in city nightlife.

dialects enabled linguists to examine a relatively pure language variety (Chambers and Trudgill 1998); this methodological choice reflected the intuitive notion that exposure to other dialects is likely to alter a speaker's linguistic output, and does not seem to have been based on any systematic study of the effects of dialect contact. However, as society is perhaps increasingly characterized by mobility and contact between speakers of different varieties, the relevance of focusing solely on the ever-endangered NORM is being called into question (e.g. Chambers 2002).

Early work on second dialect acquisition tended to focus on children and young adults, in keeping with the idea that intraspeaker linguistic changes become less likely after the critical period. Probably the best-known work along these lines is Payne (1976)'s study of the acquisition of phonological and phonetic variables by children who had moved from various dialect regions to King of Prussia, Pennsylvania. Payne found that "phonetic" features of this region's dialect (fronting of (uw) and (ow), raising of (oy)) were for the most part acquired or partially acquired by these children, while "phonological" features such as the lexically conditioned distribution of tense (æ) were not; moreover, the extent of acquisition was inversely correlated with age of arrival in King of Prussia.

Chambers (1988) and Chambers (1992) similarly examined the acquisition of several new dialect features by six Canadian-born preteens and teenagers who emigrated to Oxfordshire in southern England in the early 1980s. Chambers used reading lists and other activities to elicit tokens relevant to the following five phonological variables: flapping, low back vowel contrast, low vowel backing, (r)-lessness, and intrusive (r). Like Payne, Chambers found that acquisition of the new dialect forms varies depending on age of arrival in the new region and the phonological status of each feature.[8] For instance, Chambers' subjects showed much more loss of the native flapping rule than acquisition of either the low back distinction or vowel backing, which he analyzes as involving either acquisition of a distinction or creation of a complex lexical rule. Moreover, while flapping was suppressed to some extent in all of his speakers, only one produced [r]-less forms with any frequency. Chambers concludes from these data that it is easier for D2 learners to lose rules associated with the old dialect than to learn novel rules associated with the new variety. In addition, at least two of the variables - flap suppression and vowel backing - seem to be acquired in a lexically gradual manner.[9] Tagliamonte and Molfenter (2007) took a more in-depth look at flap suppression in their study of 3 Canadian-born children under the age of 5 who moved to York, England. The children were recorded while engaging in interactive play activities for about an hour each day of the 5 years that the family stayed in England, providing an exceptionally rich longitudinal data set. Tagliamonte and Molfenter found that all 3 children accommodated to the new dialect enough to "sound British", increasing their frequency of non-flapped stop variants over time, though their grasp of the sociolinguistic constraints governing use of these variants lagged behind their understanding of the relevant linguistic constraints. The relative ease with which the flapping rule is abandoned by North American children in the UK seems to be shared by adults. Shockey (1984) impressionistically examined the extent to which 4 native speakers of American dialects living in Essex have reduced flapping

[8] Chambers also notes that the orthographical transparency of a feature seems to play a role; if a non-native contrast is recoverable from spelling, then it may be easier to acquire than a contrast which is not represented orthographically.

[9] Chambers has the following to say about the mechanism of acquisition: "phonological innovations are actuated by the acquisition of instances of the new rule or phoneme and only become rule-governed or systematized . . . after a critical mass of instances has been acquired" (p. 663). Though he frames his discussion in terms of generative phonology, this statement well captures the spirit of the Exemplar Theoretic view of phonological change.

in their speech. She found that all four speakers suppressed flapping to a significant extent, with one speaker flapping /t/ only 17% of the time. Moreover, speakers suppressed flapping more often for /t/ than /d/, which Shockey hypothesizes is due to the greater phonetic similarity of [d] and flap.

Recent work has begun to examine the D2-acquisition capabilities of adults who are well past the critical period. The observation is commonly made that people who move from one dialect region to another seem to shift their accent towards that of their adopted region; such anecdotal perceptions were systematically investigated by Munro et al. (1999) in their study on the perception of dialect differences in Canadians who had moved to Alabama. Both lifelong Canadians and lifelong Alabamians judged speech samples from these adults as sounding different from their own dialects, indicating that some (but not complete) acquisition of D2 features had occurred.[10] While Munro et al. did not examine the extent to which particular features were adopted, a few other studies have attempted to quantify D2 acquisition in this manner. Bowie (2000) compared lifelong natives of Waldorf, Maryland to "exile" natives of Waldorf who had moved away from their hometown, to see whether the latter had changed aspects of their speech as a result of contact with a second dialect. A variety of features were examined in this study, including realization of (ey) in various phonetic environments, fronting of prenasal (æ), and the production and perception of the *pole/pull/pool* distinction(s). Bowie's results varied from speaker to speaker, leading him to observe that "the biggest generalization that can be made here is that those things that are not already in a state of change in the native dialect appear more resistant to change upon constant exposure to a new dialect than those things that are in a state of change", a

[10]It is also possible that the native Canadians may have altered their speech upon arrival in Alabama, but not towards the Alabamian speech variety.

conclusion which is consistent with the findings of Nahkola and Saanilahti (2004) and Sankoff and Blondeau (2007).

In a smaller study, Conn and Horesh (2002) looked at the speech of two adult natives of Michigan who moved to Philadelphia. Conn and Horesh focused on the realization of two variables, (ow)-fronting and (æ)-tensing, which are realized differently in these regions: Northern Cities-shifted Michigan has a relatively backed (ow) and phonetically tense (æ) in all contexts, while Philadelphia is characterized by fronted (ow) and a split between tense and lax (æ) that is subject to various phonological and lexical conditioning factors. The choice of these two features allowed Conn and Horesh to compare the acquisition of a phonemic distinction to that of a simple phonetic shift. Their results seem to indicate that this distinction does not constrain acquisition as much as we might expect, based on, for example, Payne's study. One subject did not alter her production of (æ), but did show fronting of (ow). The other speaker, however, behaved in the opposite way: he did not accommodate to Philadelphia (ow)-fronting,[11] but did show evidence of a tense-lax (æ) opposition (though the resulting pattern was a nasal system, not an accurate replication of the Philadelphia system).

Sankoff (2004) also presents a study of two speakers, in this case participants in the *Seven Up* series who were recorded every seven years from age 7 to 35. Both speakers were from northern England, but later moved to other regions: Nicholas, from a rural part of Yorkshire, went on to study at Oxford, marry a speaker of southern British English and move to the United States, while Neil, a lower middle class kid from a suburb of Liverpool, attended university briefly in Aberdeen before

[11]There may, of course, be extralinguistic reasons for this result. For at least some speakers of American English, fronted (ow) is a marker of gay speech (e.g. Munson et al. 2006); if the male speaker does not identity as gay, then his lack of accommodation to Philadelphia fronted (ow) might be due to avoidance of this marker.

moving on to a life of mostly unemployment in various parts of the UK. Sankoff looked at their use of two features which set northern varieties of English English apart from Southern varieties, Broad-A and FOOT/STRUT. In northern dialects, a shorter front [a] is used in certain TRAP class words like *last* and *chance*, while these words are realized with [ɑː] in southern varieties. Northern varieties are also characterized by a lack of contrast between the FOOT and STRUT lexical sets; words in both of these classes are produced with [ʊ] in the North, while STRUT words are realized with [ʌ] in the South. The two speakers showed somewhat different trajectories of changes for each feature. While Nicholas remained an [a] user in words like *chance*, Neil showed more use of non-northern [ɑː], especially by age 35. Meanwhile, both speakers showed some evidence of acquiring a FOOT/STRUT distinction, in that STRUT words were variably produced with unrounded variants, while FOOT words showed no hypercorrection away from [ʌ]. However, it must be noted that Neil did not have a particularly round STRUT in early life either, for unknown reasons. Considering both of these features together, these results are striking, for they indicate that in this case, the acquisition of a new contrast seems to be easier than reassigning lexical items to an already-existing category. Sankoff suggests that this is due to the relative salience of the two features to northern speakers: the salience of Broad-A as a marker of southern speech makes northerners reluctant to adopt it, while the FOOT/STRUT difference is below conscious awareness, and thus more vulnerable to shift.

More recently, Evans and Iverson (2007) examined accent change with respect to Broad-A and FOOT/STRUT in university students. About 20 native speakers of a northern, nonstandard variety of British English were recorded reading a set of words in carrier phrases at four time points: just before beginning university, 3 months later, and then again after 1 and 2 years had passed. Of particular interest in this study was the production of the words *bud, cud, could* and *bath.* While *bud, cud*, and *could* are all produced with the same [ʊ] vowel in the Northern variety natively spoken by the subjects in this study, *bud* and *cud* are produced with a [ʌ] in the Standard Southern variety encountered by these students at university. Tokens of *bath*, meanwhile, are produced with a shorter, fronter low vowel [a] in the North and [ɑː] in the South. Evans and Iverson found that the speakers in their study did change their accents over time, though these changes did not involve straightforward, entirely accurate mimicking of the standard dialect. First, speakers shifted *bud, cud*, **and** *could* to a more centralized, [ʌ] -like position, hypercorrecting FOOT-class *could* in a way that Sankoff's speakers did not. Evans and Iverson discuss a few possible reasons for this. It might be the case that speakers did not (or could not) create a new phonological category, and thus simply shifted their existing [ʊ] to a more centralized position, reflecting the phonetic distribution of the (perceived) single corresponding category in the southern dialect. Or, perhaps speakers did create a new category, but hypercorrected their [ʊ] words into it, in order to inject even more standard-sounding vowels into their speech. *Bath*, meanwhile, changed slightly, but the quality of the vowel remained [a]-like; Evans and Iverson speculate that this may be due to speakers retaining the vowel category while making small adjustments to "soften" their accent. While it is difficult to say exactly what kind of input the speakers in this study received, its findings do show that adult speakers are capable of shifting their accents in measurable ways over time.

If there is any consensus to be taken from these second dialect acquisition studies, it is that speakers can adopt features of a new ambient dialect over time, though there is much individual variability in how specific features shift. There is some indication that the linguistic status of the feature is important, but it is difficult to draw any firm conclusions about the linguistic constraints on second dialect acquisition, due to low token counts, limited lexical scope, and small speaker samples. Moreover, the effects of internal factors seem to be heavily mediated by social factors such as individual experiences, exposure, and attitudes.

1.3.3 Imitation of fine phonetic detail

Laboratory phonology studies of accommodation to new speech input have explicitly sought to examine the linguistic mechanisms underlying these shifts. These studies of the fine-grained changes that speakers make to their phonetic output were initially inspired by the general question of speech normalization: how does an infinitely variable signal get translated into an invariant underlying representation? Some researchers began to think that this was the wrong way to frame the problem, since it assumes that what speakers store is a highly abstract, redundancy-free representation that excludes all information which seems nonessential from a linguistic standpoint (e.g. Joos 1948, McClelland and Elman 1986, Morton 1969, Studdert-Kennedy 1976). However, contrary to Halle (1985)'s assertion that "we practically never remember most of the salient acoustic properties" of heard utterances, people do seem to recall such details, as indicated by both anecdotal reports and systematic studies showing that people retain memories for words spoken in particular voices (Hintzman et al. 1972, Cole et al. 1974, Mullennix et al. 1988) and with particular intonational contours (Schacter and Church 1992, Church and Schacter 1994). Since speakers can remember and make use of these linguistically-irrelevant details of the acoustic signal, it makes sense to then ask whether they are similarly affected by phonetic detail which is crucial for identifying linguistic units such as words and segments - for example, variation in vowel formant frequencies and voice onset time.

Indeed, research from the recent past indicates that speakers do have this ability to retain detailed phonetic information. Work by Nygaard and Pisoni (1998) and Norris et al. (2003) has demonstrated that hearers will quickly adjust their perceptions of phonemes as a result of exposure to talker idiosyncrasies; in fact, such adjustments are still evident 12 hours after the relevant exposure (Eisner and McQueen 2006). Such effects have also been observed for production: Goldinger (1998) showed that speakers will shift their realizations of words to sound more similar to tokens of those words as spoken by other voices. In this study, subjects were first instructed to read a series of written words (creating baseline tokens), and then participated in a shadowing task that involved immediately repeating the same set of words as tokens of each word (previously recorded by other voices) were presented. When the baseline and shadowed tokens were compared to the other-voiced targets in AXB discrimination tests, listeners judged the shadowing tokens to be more similar to the other-voiced targets, indicating that some shift in production had occurred. Similarity in this experiment was determined holistically: Goldinger's stimuli consisted of mono- and bisyllabic English words composed of various segments, and because he did not examine any particular linguistic feature, it is unclear which aspects of each utterance the talkers were imitating. Pardo (2006) obtained similar results using a map task to stimulate imitation between pairs of talkers; she found that utterances produced as repetitions of an interlocutor's speech during the map task were judged to be more similar to this person's speech than utterances produced before the task. Shockley et al. (2004) replicated Goldinger's study, but additionally focused on the specific feature of voice onset time (VOT), showing that talkers measurably lengthen their VOTs after exposure to target words with longer VOTs. It is important to note that for all of these tasks, subjects were at no point explicitly instructed to imitate or copy the target tokens: imitation (whether of recorded utterances or of tokens produced by an interlocutor) occurred spontaneously, consistent with the literature on spontaneous imitation and coordination of non-linguistic gesture and posture, e.g. Shockley et al. (2003).

These results, obtained as they are from tasks that involve shadowing of or accommodation to an immediately present stimulus, may not bear on questions of true phonological significance. Such findings may indicate that talkers are capable of keeping tokens in short term memory for the purpose of executing an immediate motor plan, but they do not show that change has occurred in either the grammar or the lexical representation of the relevant words. Goldinger and Azuma (2004), however, use a non-shadowing imitation[12] paradigm to show that exposure to relevant tokens may indeed alter the stored representations of words. In this study, 12 subjects recorded baseline tokens of 160 common English words, evenly divided into four groups based on their frequency of occurrence. One day later, the subjects returned to complete a task which involved identifying auditorily-presented training tokens from an array of words on a computer screen. The training tokens consisted of the same 160 words from the first phase of the experiment, spoken by 4 additional speakers who were not subjects in the experiment; tokens varied with respect to how often they were presented during the course of the listening task.

[12]While this method is known in the literature as the "imitation paradigm", this terminology is somewhat misleading; it may imply that subjects are explicitly trying to imitate heard tokens, which is crucially not the case.

Five days after completing this task, the subjects returned to record new tokens of the same 160 words. These test tokens were used along with the baseline and training tokens to construct trials for an AXB task, in which the X for each trial was the training token of a given word and the A/Bs were baseline and test tokens of this word produced by a subject. 300 listeners were instructed to listen to the AXB sequence and decide which of the A or B was a better imitation of the middle word. The results of this AXB test indicate that the test tokens were generally considered to be better imitations of the target tokens, with an important effect of word frequency and target exposure: greater exposure to target tokens during the Day 2 listening task increased the likelihood that the test tokens were judged to be better imitations of the target, and this effect was strongest for low frequency words. Again, subjects in this "imitation" study were not instructed to imitate the target utterances, but their test productions were still shifted as a result of exposure to no more than a dozen target exemplars encountered five days before test recording. This result has since been replicated by Nielsen (2005, 2006), who manipulated the specific feature of VOT in voiceless labial stops and showed that the imitation effect extends to productions of other words which contain initial voiceless stops. More recently, Delvaux and Soquet (2007) have demonstrated that subjects completing a verbal task will measurably shift their production of vowels towards those of a different regional dialect after exposure to recordings of the other dialect.

Such results have clear implications for the issue of second dialect acquisition: if adults who are past the critical period will spontaneously alter their productions of a word after hearing just a few instances of it spoken by an unknown disembodied voice in the lab, then we would expect adults in an adopted dialect region to similarly absorb the pronunciations of speakers around them, to whom they are presumably more motivated to accommodate.

1.4 Testing phonological theories with D2 data

Sociolinguistic studies of change over the lifespan generally, and second dialect acquisition more specifically, have shown that individual speakers are capable of making both quantitative and qualitative changes in their phonological behavior over time, given exposure to new input. However, it is difficult to draw any more detailed conclusions regarding the linguistic constraints on or mechanisms underlying such changes from this body of work. In the case of longitudinal studies based on found corpora (e.g. Sankoff 2004 or Bowie 2009), this is mostly because the researcher has to make do with the dataset as is; the type, quality, or amount of data contained therein may or may not be enough for or relevant to testing specific hypotheses about phonological change. However, for most sociolinguistic studies, it seems that this sparse data problem is an issue that arises not from the actual data, but from the way that this data is sampled and analyzed. These choices can be attributed directly to a set of linguistic assumptions. For example, if the linguist assumes that change in a particular vowel category equally affects all words containing that vowel, then she may be content to collect ten tokens of the category to establish its realization in phonetic space. Unfortunately, ten tokens will not be enough to allow later testing of the idea that there are word-specific effects on phonetic variation.

In contrast, lab experiments of the type reviewed in section 1.3.3 are especially set up to test linguistic theories by examining fine-grained phonetic detail, phonological constraints affecting phonetic realizations, and the effect of lexical frequency. However, the conditions under which these experiments are carried out, while precisely controlled in terms of linguistic factors, are rather artificial, and it is unclear to what extent conclusions based on lab behavior can be generalized to real-life contact situations. Moreover, such studies can usually only track very short term accommodation.

The current study is intended as a third way between these two approaches. It is a sociolinguistic study of second dialect acquisition, examining the linguistic behavior of born-and-bred Canadians who moved as adults to the New York City region. However, the analysis is constructed with the primary intent of testing different theories of representation. The population of speakers was chosen specifically because their native dialect and adopted region's dialect differ with respect to phonological features that may be especially relevant to deciding between these theories, and the data was gathered and analysis carried out so that the linguistic questions of interest could be answerable.

Chapter 2 of this dissertation describes the methodology of the study. Chapter 3 details the phonological predictions made by both generative and usage-based theories regarding the type of behavior that speakers in the study should exhibit. Chapters 4, 5, and 6 present the linguistic analysis and results. Chapter 7 presents findings on the effects of various social factors on the data. Chapter 8 contains conclusions and questions for further research.

CHAPTER 2

METHODS

2.0 The interview study

How does a speaker's phonological system change as as result of long-term exposure to new dialect input?

In an ideal world, one would approach this question methodologically by designing a longitudinal study. A group of speakers living in some dialect region of interest would be selected and their speech recorded; the group would then move to some other area characterized by a different phonological system, and their speech rerecorded after some amount of time spent immersed in this new dialect. From the observed changes in linguistic behavior, the linguist would then infer, where appropriate, changes in underlying representations or phonological processes. In the real world, achieving this level of control in a longitudinal study of mobile people is unlikely, given the difficulty of finding speakers who have the native system required, are moving to the/a region of interest, *and* are willing to participate in a study over the long term. Financial and temporal limitations can also make such studies impractical.

One way around these problems is to rely on fortuitously available longitudinal speech data that was not necessarily collected for purposes of linguistic study; notable examples include Sankoff (2004)'s use of data from the 7-Up television series and Harrington (2006)'s study of the Queen of England's Christmas broadcasts, as described in the previous chapter. Of course, in order to conduct a post hoc longitudinal study of this type, one must be lucky enough to discover a corpus that is amenable to linguistic analysis, whether auditory or acoustic. However, corpus relevance is perhaps a more serious concern: because the linguist has no control over the style of speech collected, the speakers recorded, or the types (and token counts) of linguistic features elicited, she is limited in terms of the questions she can ask and answer. Intriguing lines of inquiry may be suggested by the particular data set available, but these might be entirely different from the theoretical questions of primary interest. Returning to the question posed at the start of this chapter, there seems to be no readily-available corpus of speech gathered from individuals who have moved from one dialect region to another; I might have searched various sources of online recordings (for example, news programs) and turned up a few serendipitous recordings of mobile speakers, but it is unlikely that these would have been in any way comparable, or contain (enough of) the relevant data to answer the specific theoretical questions raised by this study.

If one is not able to gather theoretically relevant longitudinal data in either a planned or post hoc manner, the next best strategy is to find and study speakers in the present whose linguistic past can be determined with reasonable confidence. This is the procedure followed by Conn and Horesh (2002) and Bowie (2000), two studies which sought to determine how mobile speakers' language had changed after moving from the native region to a new dialect region. Both of these studies analyzed speakers at one point in time, after they had been living in their new region. However, in each case the speakers' likely initial system could be established based on independent knowledge about the dialect of the native region. The researchers could then infer what phonetic and phonological changes their speakers had undergone as a result of relocating.

This is the approach I took to studying the long-term effects of new dialect exposure. In the remainder of the chapter, I describe the group of speakers I chose to work with - born-and-bred Canadians who moved as adults to the New York/New Jersey region - as well as the linguistic characteristics of their native and adopted regions which make this group particularly apt for study. I also discuss the structure of the sociolinguistic interview that each speaker participated in, as well as other details of data collection.

2.1 Why Canadians in the NYC region?

Expatriate Canadians in the New York region were an ideal group to target for this research for at least two reasons. First, Canadian English differs from the varieties spoken in and around New York in several documented and measurable features (described in more detail below). Second, for born-and-bred Canadians whose families have been in Canada for at least a few generations, I could with reasonable confidence infer the state of their dialects prior to emigrating to the United States, at least with respect to the phonological features examined in this study.

2.1.1 Canadian English

Boberg (2008) states that, with the exception of the variety spoken in Newfoundland, "Canadian English is remarkably homogenous from one end of the country to the other" (p. 146). Moreover, Canadian English (henceforth CE) shares many phonological features with other North American dialects of English. In consonantal inventory and patterning, CE is virtually indistinguishable from varieties spoken below the 49th parallel: post-vocalic (r) is retained, trochee-medial alveolar stops are flapped, and younger speakers tend not to produce the historically realized yods in words like *news* and *tube* (Labov et al. 2006, Boberg 2008). In the vowel space, CE has much in common with the English spoken in the western part of the United States; indeed, Labov (1991) grouped the American West and Canada together as the third major dialect region of North American English. Like the West, Canada does not participate in the Northern and Southern vowel shifts, and is characterized by the unconditioned merger of the low back vowels in words like *cot* and *caught* (see Section 2.1.3) as well as by the merger of nonhigh front vowels before /r/, such that the words *Mary*, *merry*, and *marry* are produced homophonously.[1]

However, there are several features which distinguish CE from American dialects. The Canadian Shift involves the backward movement of (æ) and subsequent lowering and backing of (ɛ), then (ɪ); first noted by Clarke et al. (1995) in Ontario, it has also been found in Montreal (Boberg 2005) and elsewhere throughout Canada (Labov et al. 2006), except for the Maritime Provinces. An especially salient feature of CE is Canadian Raising, in which the diphthongs (aw) and (ay) are realized with higher nuclei before voiceless obstruents; while raising of (ay)

[1]The exception is the dialect spoken in Montreal, which merges *Mary* and *merry*, but keeps *marry* distinct.

does occur in several American dialects, raising of (aw) is more restricted to CE (see Section 2.1.4). Another noted feature is the production of pre-rhotic low back vowels in words like *sorry* and *tomorrow*, with *sorry* [sɔri] in particular serving as a shibboleth of Canadian speech (vs. [sɑri], which is more typical of U.S. dialects). Finally, Canadians may be setting themselves apart from American speakers by their treatment of so-called foreign-a words, which arguably involves the creation of a new phoneme (Boberg 2000).

The 17 speakers interviews in this study were all born and raised within the homogenous CE region described above: crucially, a region characterized by a merger between the low back vowels and raising in pre-voiceless (aw).

2.1.2 New York-area English

The English spoken in and around New York City has been a point of sociolinguistic interest since Labov's landmark study (Labov 1966). Much of the focus here has been on consonantal variables, notably the realization of post-vocalic (r) and the stopping of interdentals in words like *these* and *those*. However, ample attention has also been paid to vocalic features. New York City English exhibits a robust contrast between the low back vowels of *cot* and *caught*, though the quality of the vowel in *caught* varies according to class and style (Labov 1966) as well as age and ethnicity (Becker 2010); this distinction is shared by adjacent areas of New York state, including southern Westchester County and the western half of Long Island, as well as neighboring areas of Northern and Central New Jersey (Labov et al. 2006). New York City is also characterized by a complex allophonic conditioning of the realization of (æ), such that the (æ) before tautosyllabic /m/ and /n/, voiceless fricatives, and voiced stops is produced with a tenser quality than (æ) in words like *pat* (Cohen 1970, Labov 1972, 1994).[2] The situation regarding (æ) in nearby areas of New Jersey is more complex; some speakers exhibit a NYC-like tense (æ) alternation, while others use a tense allophone only before nasals (Ash 2002, De Decker and Nycz 2006).

Fourteen of the speakers in this study live within New York City (specifically, Manhattan, Queens, or Brooklyn). The remaining three speakers live in neighboring counties of New Jersey (Essex, Middlesex, and Monmouth counties), all of which are located within the New York City metropolitan area (Office of Management and Budget 2009). The English spoken in these areas of New Jersey differ in various ways from that spoken in New York City; New Jersey varieties are traditionally (r)-full, for example, and exhibit somewhat different (æ)-tensing patterns as described above. Indeed, New York City and the Mid-Atlantic States are considered by The Atlas of North American English (henceforth ANAE) to belong to separate dialect regions. However, New York City and neighboring counties in New Jersey do share many linguistic features, including (crucially) the two features examined in this study: a low back vowel distinction and a lack of raising in prevoiceless (aw). For present purposes, then, I will consider all speakers to have been exposed to essentially the same second dialect, a variety which for efficiency of discussion I will reify as "New York area English" (henceforth, NYAE)

2.1.3 The low back vowel system

A key feature defining and distinguishing North American dialects of American English is the status of the low back vowels. The feature is special among phonological variables of North American varieties because it involves variation in phonemic in-

[2]Younger speakers and ethnic minorities, however, do not seem to be showing this complex conditioning (Becker and Wong 2009).

ventory: speakers of dialects which distinguish words like *cot* and *caught* have two low back vowel phonemes, /ɑ/ and /ɔ/, where speakers of dialects exhibiting the merger of these two words classes have only one phoneme, /ɑ/.

In many dialects of American English, this feature seems to be in flux: the merger can be variably present across (and sometimes within) speakers within a community. However, in the two dialect regions that are the focus of this study, the situation is more clear-cut. While data from Canada in ANAE is sparse, Canada is included within the larger merged region (Labov et al. 2006). Moreover, according to Boberg (2008),"virtually all native speakers of Canada today" have this merger, which has been present in Canadian English for several generations. It is therefore reasonable to assume that speakers acquiring their native dialect of English in Canada are acquiring a variety which does not distinguish two low back vowels.

The situation in New York City and surrounding areas is quite different. The mid-Atlantic region is noted in ANAE as being one of a few areas in which the *cot/caught* distinction remains robust; here, the raised quality of the vowel in *caught* protects the contrast (though recent work by Becker (2010) indicates that the phonetic distance between these vowels is decreasing over time). In neighboring New Jersey, these vowels are also distinct. Coye (2009) reports, based on questionnaire data, that the merger of these vowels is "gaining a solid foothold in New Jersey"; this is perhaps true of counties in the northwest, where around 30% of questionnaire respondents report that the vowels of *Don* and *Dawn* sound the same, but the responses for the majority of counties in NJ are overwhelmingly (>85%) distinct.

2.1.4 Canadian Raising

Perhaps the most salient feature of CE is Canadian Raising, the raising of the nuclei of words in (aw) and (ay) before voiceless consonants (Joos 1942). Canadian Raising is most commonly analyzed as being the result of a phonological rule (e.g. Chambers 1973), as the quality of the nucleus is fully predictable from context.

Despite its name, Canadian Raising is not limited to Canada. Raising of (aw) has also been documented in Virginia (Kurath and McDavid 1961) and Martha's Vineyard (Labov 1963). (ay)-raising is even more widespread in American dialects, again in Martha's Vineyard, but also Philadelphia (Labov 1994), the Inland North (Eckert 2000), and Ocracoke Island (Schilling-Estes 1998). However, this feature, at least with respect to (aw), is still largely associated with CE; the phrase *out and about*, produced with hyper-raised nuclei ("oot and aboot") is a popular, if phonetically inaccurate, stereotype of CE.

The New York City region, in contrast, does not exhibit raising of either (ay) or (aw). Labov et al. (2006) notes the "conservative character of New York City upgliding vowels", in particular the facts that the nuclei of (ay) and (aw) are no higher than those of the low vowels (æ) and (o).

While both the low back vowel system and prevoiceless (aw) vary with respect to their realizations in many North American dialects, the status of each of these features in Canada and the New York area is more certain. An individual who acquires CE as his first dialect will develop both a low back vowel merger and some degree of Canadian Raising, especially in (aw). If this speaker then moves to the New York metropolitan region, he will encounter a low back vowel distinction and a lack of Canadian Raising.

2.2 Recruitment and selection of speakers

Interviews were conducted in New York City and neighboring counties in New Jersey from Fall 2006 through Summer of 2008, with 17 native Canadians who had moved to the New York metropolitan region as adults.[3] Canadians in this region do not cluster in particular blocks or neighborhoods; there is no Canadatown or Little Ontario in which the linguistic fieldworker can search for potential interviewees. Thus, speakers were recruited almost entirely via online fora. These included a facebook group ("Canadians in New York"), a Canadians-in-New-York MeetUp.com group, and Connect2Canada.com, a website sponsored by the Canadian Consulate which allows expatriates in the U.S. to find other Canadians in their area. I also met several speakers at a Central Park hockey game for alumni of four Canadian universities, after contacting the McGill-Alumni-in-New-York association.

In online posts and in-person descriptions, I explained that I was interested in learning about peoples' experiences moving to and living in a new region, particularly with regard to cultural and communication differences. There was no hiding the language-centric nature of my research, given my Linguistics Department affiliation. By expressing an interest in "communication" generally, I hoped to move the interview focus from accent to more discourse-level attributes of the relevant dialects, to avoid too much direct discussion of the phonological features under study; to some extent, this worked. However, the unavoidable salience of

[3]All speakers relocated to to the area after age 21.

pronunciation differences possibly helped my recruitment efforts, as speakers often mentioned an avid interest in accent differences in their email responses to me. I also stated that I was particularly interested in people who had positive attitudes about and experiences with living in the NY/NJ region and interacting with "the natives."

The nature of my sample differs from typical sociolinguistic studies in two important respects. First, it is important to stress that this is not a community study. As noted in the previous paragraph, there is no geographically-defined group of Canadians in New York, and while there may very well be smaller networks of expatriate Canadians in the region,[4] my own collection of speakers exhibits no such cohesiveness: none of the individuals I interviewed knows any of the others, and all vary widely in terms of their province of origin, experience coming to the U.S., local social network, and other factors. Second, this is by no means a socially balanced sample. Ideally, the group of speakers in this study would be balanced for gender, age, time spent in the United States, place of origin, and other surely relevant social factors, but ultimately I had to be satisfied with whoever responded to my posts and had time to be interviewed. In addition, all of the speakers have had at least some university-level education, and can be described as middle or upper middle class. I do not think these are serious limitations for the current study; my aim is not to make claims about the behavior of "Expatriate Canadians in New York" per se, but to examine in fine detail the types of linguistic changes that occur when speakers characterized by a particular set of linguistic properties are exposed to a particular new set of properties.

[4] Communities of Hockey Practice?

Speaker	Gender	Age	Years in NY/NJ	From
LC	female	30	1	Ottawa/Toronto
LW	female	31	10	New Brunswick
PW	male	32	<1	Vancouver/Toronto
BW	male	37	2	Toronto
NW	female	39	14	Alberta
TM	female	41	3	Toronto/Manitoba/Ottawa
ES	male	42	5	Manitoba/Alberta
JF	female	45	14	Manitoba
LG	female	46	7	northern Ontario/Toronto
JC	male	48	18	Montreal
EW	male	50	16	Saskatchewan
BK	female	54	21	Ottawa/Montreal
GH	male	54	15	Montreal/Toronto
CW	female	54	28	Montreal
SS	female	54	27	Montreal
DB	female	58	11	Halifax/Toronto
VJ	female	70	44	Toronto

Table 2.1: Speakers in the study, described by gender, age and number of years spent in NY/NJ at time of interview, and region of origin.

Table 2.1 summarizes some basic relevant information about the 17 speakers: gender, age at time of interview, the number of years spent in the New York region at time of interview, and region of origin. These and other social characteristics will be examined for their possible effects on linguistic behavior in Chapter 7.

2.3 Recording equipment

All interviews were recorded directly to 16 bit, 44.4Hz WAVfiles with an Edirol (by Roland) R-09 digital recorder, using an electret condensor lapel mic.

2.4 Interviews

Each speaker participated in an in-person sociolinguistic interview. Interviews were about an hour and a half in length, and took place either at the speaker's home, at my office in the Linguistics Department at New York University, or in a public restaurant or café. The discussions began with basic questions about the speaker's background and where they grew up in Canada, later moving to their reasons for coming to the United States and their experience doing so. Large parts of the interview often focused on the speaker's career, because that was usually a salient reason for moving. I also elicited speakers' opinions of the area where they grew up and their adopted region, and encouraged them to compare their new and old homes at both a local and national level (e.g. Toronto vs. New York City, Canada vs. the U.S). After about an hour of conversation, we moved to reading and judgment tasks (described in more detail below). After these tasks, the interviews resumed with discussion of language and accent issues, which everyone had much to say about.

2.5 Reading tasks

Each speaker completed 2 reading tasks. First, speakers were asked to read words presented in isolation on flashcards. Next, speakers were asked to read and provide judgments on a list of minimal or rhyming pairs.

2.5.1 Word list

Speakers were asked to read out loud 135 words which were presented on flashcards. These words represented a variety of word classes, though low back vowel words

awed	*bought*	*caller*	*caught*	*dawn*	*dog*
odd	pot	collar	cot	don	cog
naughty	*pawned*	*talk*	*tall*	*walk*	
knotty	pond	sock	doll	wok	
law	*paw*				
la	pa				
bother	logger	*daughter*			
father	lager	*water*			

Table 2.2: Low back vowel pairs elicited in the Minimal & Rhyming Pair task. Italicized items are /ɔ/-words in NYaE, and non-italicized items are /ɑ/-words

featured prominently in the list. Many of these low back vowel words were also present in the minimal pair list, enabling a comparison of vowel production across styles. Two versions of the word list were used over the course of data collection. The original word list (presented to the first 5 speakers interviewed) included fewer low back vowel words; once it became apparent that there were differences in how these vowels were produced across word list and minimal pair styles, more of the minimal pair list words were added to the word list to enable a more robust comparison across contexts (more details on this are provided in chapter 4). The complete word list can be found in Appendix A.

2.5.2 Minimal & rhyming pairs

Speakers also completed a sociolinguistic minimal pair task and a rhyming pair task. Each speaker was handed a printed list of minimal pairs, and asked to read each pair out loud, then say whether the pair sounded the same or different. Speakers were also given a shorter list of rhyming pairs and asked to pronounce each pair, then say whether the pair rhymed. Each of these lists primarily probed the low back vowel distinction, though these pairs were interspersed with other pairs of potential interest (e.g. *Mary/merry*). A list of the low back vowel pairs elicited is shown in Table 2.2; the complete list of minimal pairs and rhyming pairs can be found in Appendix B.

2.6 Judgment task

After completing the canonical minimal and rhyming pair task, speakers were then asked to look back over these two lists and say whether they thought people from New York region would either have difference judgments of some of these pairs, or pronounce particular words differently. The purpose of this task was to assess whether speakers have a conscious awareness of the low back vowel distinction in NYaE. Speakers were encouraged to produce these forms "as a local would say them", so that I might get a better sense of what they believed the local phonetic targets for relevant words to be.

The next chapter will discuss in more detail the phonological theories introduced in Chapter 1, and the precise predictions these theories make with respect to how the speakers in this study should alter their realizations of the low back vowel system and (aw).

CHAPTER 3

PHONOLOGICAL REPRESENTATION AND CHANGE

3.0 Two theories of underlying representation

The mental representation of words and sounds is one of the central concerns of phonological theory. While there are many specific theories regarding the nature of these representations and how they map onto surface forms, these diverse views can generally be subsumed under two major groups: *generative models*, in which underlying representations are highly abstract and far removed from observed surface forms, and *usage-based models* which claim that these representations contain much phonetic detail, reflecting actual heard tokens.

In this chapter, I will begin with an overview of both types of models, paying particular attention to how they deal with two major concepts: *contrast* and *phonological generalizations*. I will then describe the mechanisms underlying intraspeaker variation and change in each of these models, as well as their implications for how speakers might acquire a new contrast or alter a generalization after exposure to appropriate linguistic input. Along the way, I will set out the specific predictions each theory makes regarding how Canadian English (CE)speakers should acquire features of New York City area English NYaE, so as to lay the foundation for the linguistic analyses in Chapters 4, 5, and 6.

3.1 Representations in generative theory

The generative view, which is also the mainstream view in phonology, is that underlying representations are quite abstract compared to surface forms. This view of representation has a long history in phonological thought, being a principle component of Structuralism (de Saussure 1916) and the linguistic theories of the Prague School (Trubetzkoy 1969[1939], Jakobson 1962). However, it is most closely identified in the modern linguistic era with The Sound Pattern of English (Chomsky and Halle 1968). Later developments within generative theory such as autosegmental phonology (Goldsmith 1979) and feature geometry (Clements 1985, Clements and Hume 1995) made the underlying representations more complex, but continued to hew to the same principle of abstractness from surface form.[1]

As with most ideas in linguistics, this notion of abstractness comes bundled in a larger theoretical package of interconnecting assumptions and principles. For this reason it is useful to consider various key components of the larger framework; indeed, doing so is crucial for understanding the types of predictions this theory makes with respect to language change over the lifespan.

[1]The reader will note that this chapter does not compare rule-based theories of phonology with Optimality Theory. This is because the question of rules vs. constraints is orthogonal to the issue of representation: an OT analysis which acts upon an abstract, feature-based lexical representation will ultimately predict the same sort of results with respect to the likelihood of intraspeaker change as the rule-based model described here.

Representations are contrastive and minimally-specified. According to Chomsky and Halle (1968), the phonological portion of each lexical entry consists of a sequence of feature matrices, with each matrix corresponding to a distinctive phoneme. SPE-type features are typically given phonetically inspired names such as [VOICE] or [NASAL], but these labels are essentially mnemonic, as the real purpose of features is to distinguish matrices (and potentially words) from one another. The articulatory and/or acoustic spelling out of these labels is determined by phonetic implementation rules later in the derivation of surface forms; none of this phonetic information is present in the underlying representation.[2] The details concerning which features may be considered redundant and how they are filled in at later points in the phonological derivation are not entirely agreed upon (see e.g. Keating (1988) for a theory of underspecification in which some features remain underspecified throughout the derivation). However, all generative theories share the general assumption that lexical representations should contain the minimum number of features necessary to establish contrast.

Each lexical item has a single, stable representation. Closely intertwined with this idea of minimally specified underlying representations is the assumption that there is only one such representation per lexical item, from which all surface variation derives. This position is the logical endpoint of an approach which seeks to abstract away from surface differences at the segment level; it is also motivated by the need to to define the morpheme as a unit consisting of a pairing between sound and meaning while maintaining the unity of morphologically related surface

[2]The precise spelling-out of a given feature may involve a variety of phonetic properties. For example, a [VOICE] contrast may be signaled by vocal cord vibration (voicing), voice onset time, preceding vowel length, presence of prenasalization, etc. depending on context; moreover, it seems that no particular phonetic property - including literal voicing! - is a necessary condition for realizing the [VOICE] contrast.

alternants (e.g. *knife/knives*, the plural forms [-s], [-z], [əz]) (Harris 1942). Because these representations do not reflect surface variation, they are stable over time; though the underlying representation for a morpheme can in principle change, via the addition, subtraction, or alteration of one or more features, this is assumed to be a rare occurrence within the system of an individual speaker.

Representations are separate from generalizations. As noted above, these unique, minimally-specified underlying representations serve as the input to phonological rules; these rules alter features of the representation to produce intermediate and ultimately surface forms, whose actual realization is determined by implementation rules at the very end of the derivation. *Representations* and *rules* are distinct components in this theory: underlying representations contain all - *and only*- that which is arbitrary and unpredictable about the word form (such as segment order and contrastive features), while rules capture broader generalizations which apply to sounds in particular contexts across word forms.

No word-specific phonetics. As a result of this derivational relationship between minimally specified representations and detail-imbuing rules, words do not, and cannot, "have their own history". Rules affecting a given segment in a particular context apply to all instances of that segment across the lexicon; because of this, there can be no systematic gradient variation between words per se, or gradual phonetic shift that affects some words and not others on a lexically unpredictable basis.[3] In generative theory, a word's surface realization is essentially the predictable phonetic sum of its parts.

3.2 Representations in usage-based theories

Directly opposed to the generative view of representation is the idea that stored knowledge of a particular word consists not of a minimal, abstract sequence of symbolic elements, but a large collection of phonetically rich memories of particular tokens of that word.[4] The most fully articulated version of this theory in linguistics is Exemplar Theory (Johnson 1997, Pierrehumbert 2001, 2002, 2003, Wedel 2004, 2006) though this idea is also central to Bybee (2001)'s theory of usage-based phonology, and has precursors in the Memory Trace Models described by e.g. Hintzman (1986) and Goldinger (1998), and the "memory images" posited by Paul (1889). Again, it is helpful to tease apart various related components of this theory.

[3]There are, of course, exceptions to phonological rules, and work within the generative framework has attempted to account for these exceptions in two general ways. Some scholars have proposed that lexical entries may be marked with exception features indicating that a particular rule or constraint should or should not apply to a particular item; this is the approach taken by Chomsky and Halle (1968) in rules-based phonology, and by Pater (2009) in OT. Guy (2007), in contrast, has used consonant deletion data to argue that exceptional forms ought to be directly represented in the relevant lexical entries. While both types of approach attempt to account for word-specific behavior, it is assumed that such behavior is indeed exceptional, and not characteristic of the entire lexicon.

[4]One simple way to summarize the conflict between generative theory and usage-based theory is in terms of how much "phonetics" each is allowed into the lexicon. However, this way of talking about representation presupposes the traditional distinction between phonology-as-linguistically-interesting-phenomena on the one hand and phonetics-as-mechanical/implementational-phenomena on the other. When framed this way, the generative position appears to be the default hypothesis, while the usage-based model is the theoretically extravagant proposal that needs to be justified. However, the generative model is far from a null hypothesis. As Vaux points out, an actual null theory of lexical representation would posit that words are stored in their surface phonetic form, since such a theory "assumes that language learners do not tinker in any way with the primary linguistic data to which they are exposed" (Vaux 2003, p. 92).

Representations are phonetically-detailed. In usage-based theories, the mental representation of lexical items reflect much of the phonetic detail of actual surface forms; in fact, they are often[5] considered to be memories of heard words embedded within the parametric phonetic space, a "quantitative map of the acoustic and articulatory space" (Pierrehumbert 2003, p. 179). Categories such as words, phonemes, and allophones are abstractions over this phonetic space: word forms correspond to clouds of remembered tokens associated with a given semantic label (e.g. DOG), and sound categories such as phonemes and allophones emerge as distributional peaks within this phonetic space. This proposal - a claim of lexical *maximality* which is diametrically opposed to the generative position outlined above - is motivated by several phenomena which have been observed in the laboratory phonology and sociolinguistic literature: language-specific phonetic implementation rules (e.g. differences in pre-fricative voicing in French and English as shown by Flege and Hillenbrand 1986, lexically specific (often frequency-related) phonetic change (Phillips 1984), and phonetic shifts (see Pierrehumbert 2003 for a thorough discussion of this background). In addition, the bidirectional influence of L1 and L2 on the realization of shared categories (as observed by e.g. Flege 1987 and Sancier and Fowler 1997) supports a theory in which representations are composed of heard instances of those categories.[6]

[5]Not all representational models which incorporate phonetic detail claim that episodic memories are the basis of these representations. Kirov and Gafos (2007), for example, present a model in which underlying representations consist of dynamic gestural scores which are updated over time.

[6]Flege (1987) accounts for this two-way influence in terms of *equivalence classification*, the "cognitive mechanism which permits humans to perceive constant categories in the face of the inherent sensory variability found in the many physical exemplars which may instantiate a category (1987: 49)"; Best (1995)'s Perceptual Assimilation Model similarly describes how L1-L2 interference can be understood in terms of the way that foreign categories are mapped onto L2 learners' native categories. Both of these models are compatible with an Exemplar Theoretic model of representation.

Each lexical item is associated with a dynamic set of many representations. Because each heard token of a lexical item is stored and tagged with a label indexing it to that lexical item, there are potentially hundreds or thousands of representations associated with each word. Usage-based theories differ with respect to how many memories are retained, and for how long (recent versions of Exemplar Theory, for instance, contain a decay parameter which allows older exemplars to be forgotten over time, e.g. Pierrehumbert 2006). In all such theories, however, the number of representations will vary depending on how often the word is encountered in speech, with more frequent words having more stored memories.

Generalizations arise from representations. A common characteristic of usage-based models is that they may make no clear distinction between representations and rules (Langacker 1987, 2000, Bybee 2001).[7] Phonological generalizations are not formalized as processes that representations undergo, but arise from the regularities present across representations. Another way to state this is to say that there is no derivationally based distinction between phonemes (qua the components of underlying representation) and allophones (qua the results of phonological rules); both types of categories are represented by clouds of tokens which form distributional peaks in the parametric phonetic space, with clouds corresponding to classical allophones being more circumscribed within this space (and presumably more evenly populated) than those corresponding to the higher-level, classical phoneme.

Word-specific phonetics. In usage-based models, every word does in fact have its own history, reflecting the assumption that lexical representations are dynamic

[7]The presence of emergent generalizations does not necessarily preclude the existence of separate generalizations or rules in the same model (see Section 3.3). However, such generalizations are typically not formalized, and remain secondary to the representations which give rise to them.

and affected by usage. Representations are continually being updated with new heard tokens, but the frequency of updating varies across lexical items, such that frequently heard items will be updated more often than rarely heard items. One result of this is that frequent items might be expected to reflect ongoing phonetic changes in the broader community before less frequent items.

3.3 Hybrid models

A third type of model which combines features of both the generative and usage-based theories is also possible. Pierrehumbert (2006) endorses just such a hybrid model, noting that generative and usage-based models have each been developed to account for particular phenomena which are not easily explained by the other type of theory. Similarly, Goldinger (2007) argues for a complementary-systems approach which incorporates both stable abstract representations and dynamic episodic memories; this approach is proposed to have a neural basis, in the reciprocal interactions between a "hippocampal" network which quickly stores memories of particular events and a "cortical" network which slowly abstracts knowledge from surface statistical regularities.

Pierrehumbert (2006) convincingly claims that "the future lies with hybrid models", which promise to offer a more complete account of phonological as well as sociolinguistic phenomena. However, the distinction between generative and usage-based models drawn in the previous two sections remains a useful one, as each, in some form, will likely constitute a component of any hybrid model which is ultimately adopted. An important part of the development of such a model will be defining the particular role each component plays in various phonological phenomena: for instance, to what extent can specific instances of intraspeaker change be localized within one part of the hybrid system? To this end, I will continue to separate and contrast the generative and usage-based approaches in the discussion that follow, bearing in mind that, ultimately, a complete account of phonological representation will incorporate aspects of both.

3.4 Phonological contrast

The issue of contrast "lies at the doorstep of phonemic theory" (Goldsmith 1995), and has been a central concern of phonological thought throughout the twentieth century (Anderson 1985). It is therefore an obvious point on which to compare the two types of theories outlined above. What does it mean for two categories to contrast in each of these theories? How is contrast represented in the lexicon?

For generative phonology, stating that two sounds *contrast* means that the segments "[differ] in at least one feature" (see Chomsky and Halle 1968, p. 336 for a formal definition). Contrast in this framework is a binary notion, such that the phonological representations of two sounds/words either contrast (because they differ in one or more features) or are identical. The putative categorical nature of contrast is reflected in native speaker judgments of minimal pairs: two strings that differ in only one sound are either "the same word" or "different words", never "*sort of* the same word".

Usage-based theories, on the other hand, seem to allow for a more gradient notion of contrast. Two clouds in the continuous phonetic space, each corresponding to a different sound category, may 1) completely overlap, 2) be completely separate, or 3) partially overlap. 1) clearly corresponds to the generative case in which two sounds are identical; 2) corresponds to a generative contrast, but what about 3)? This case is potentially complicated, given the dual nature of contrast.

Deciding whether two sounds contrast is a matter of finding out whether the difference in sound makes a difference in meaning, as determined by native speaker intuitions. However, this criterion collapses the production and perception aspects of contrast. If the overlap of two categories is slight enough such that the items in question could be distinguished by listeners most of the time, then this would probably be considered a case of contrast. However, if the overlap is larger, such that the two categories were still distinguished in production but could not be identified by listeners at some arbitrary level of reliability, then the linguist might decide that there is no contrast.[8] Where exactly does one draw the line? One possibility is that there is no real line to draw. Contrast may be a gradient notion: a statement of how reliable a particular phonetic difference is for identifying a category, such that two sounds could be more or less contrastive (Scobbie and Stuart-Smith 2008).

In fact, the situation in usage-based theory with respect to contrast is even more complicated than this. Contrasts may not only be phonetically gradient in the way described above, but lexically gradient as well: because different words containing the "same" vowel category may occupy somewhat different places in the parametric phonetic space, different words (or, more to the point, potentially homophonous word pairs) may show greater or lesser contrast (e.g. *taught-tot*

[8]This last scenario describes the now well-attested phenomenon of incomplete neutralization. For example, in languages which are typically described as having a phonological rule of word-final devoicing, phonetic analysis of final stops shows that, in fact, the "voiceless" surface obstruents which are underlyingly [+VOICE] are phonetically more voiced than surface voiceless obstruents which have always been [-VOICE] (e.g. Giannini and Cinque 1978, Port et al. 1981, O'Dell and Port 1983, Dinnsen and Charles-Luce 1984, Port and Crawford 1989). Moreover, listeners can identify potentially ambiguous words based on this contrast - but only about 70-80 percent of the time (Port and Crawford 1989). Incomplete neutralization phenomena resemble the near-mergers which have been observed in studies of language variation and change (to be discussed in more detail below), and present many of the same problems for a standard feed-forward model of phonology (Nycz 2005).

might show less separation than *caught-cot*). Indeed, this kind of phonetically and lexically gradient contrast is expected in a usage-based theory.

What is the best way to go about deciding between the categorical and gradient views of contrast? The minimal pair task, which forces one of two categorical judgments, is ill-suited to investigating phenomena which are potentially gradient. Of course, it is also possible to supplement minimal pair judgments with production and perception tasks: acoustic analysis of speech can reveal whether a speaker produces a significant distinction between two potentially contrasting items, while perception experiments will show whether the same speaker is able to use this contrast to identify different words. Combinations of these tasks have revealed the existence of so-called near-mergers, cases in which speakers produce a difference which they cannot perceive (e.g. Labov et al. 1972 on *source/sauce* in New York City; Di Paolo 1988 on *full/fool* in Utah; Herold 1990 on *cot/caught* in Pennsylvania; Trudgill 1974 for *toe/too* and *bear/beer* in Norwich). These near-mergers, which are characterized by a gradient approximation of two categories, may seem to be a problem for the traditional view, but in fact they can be squeezed into a generative account: the speaker who nearly merges two phonemes is analyzed as having contrasting underlying representations identical to those of the speaker who produces the two sounds quite distinctly, but the first contrast is obscured by later low-level rules which shift the realizations of the two categories towards one another (e.g. Goldsmith 1995). Usage-based phonology is also able to account for these results, and in a way that more satisfyingly captures the difference between nearly-merged and fully-distinct speakers. However, the point is that the synchronic data yielded by one-time minimal pair judgments and production/perception tasks are going to be at best consistent with both theories of representation, at worst biased towards the categorical view of contrast and representation, and thus not useful for deciding between these alternatives.

One way to approach the problem of how to decide between alternative views of contrast is to examine the ways in which categories - and contrasts between these categories - change over the lifespan. Generative phonologists typically are not the ones who look at these data, which usually fall in the domain of sociolinguistics. However, intraspeaker change data is highly relevant to questions of phonological theory: the specific kind of changes that can occur over time in the system of an individual, and the manner in which they unfold, will depend on the nature of the representations and processes that are undergoing change. If it is possible to show that the two theories of representation described here make different predictions with respect to how contrasts may be acquired, retained, lost, or neutralized by individual speakers over time, then the relevant data can be sought out and used as evidence for one theory or the other.

In the case of the current study, the relevant data will come from speakers who have grown up with a single low back vowel category encompassing both *cot* and *caught*, then have moved to and spent time in a region which is characterized by a contrast between these two word classes. In the remainder of this section, I outline the predictions each theory makes about how these speakers should go about acquiring this contrast.

3.4.1 Contrast acquisition in generative phonology

There is little in the generative phonology literature that addresses the issue of intraspeaker linguistic change beyond the age of L1 acquisition.[9] References

[9] This is not as true in the OT literature, where several learning algorithms have been proposed and tested (see e.g. Tesar and Smolensky (1998), Boersma and Hayes (2001), Soderstrom et al.

to change exist, but these often refer specifically to cases of diachronic change (e.g. Kiparsky 1988) or are ambiguous as to whether they involve diachronic or synchronic changes.[10] However, we can speculate about the possibilities for intraspeaker change in a generative framework by extrapolating from what has been said about diachronic change within this model.[11] The most detailed work on this count has been presented within the framework of Lexical Phonology (e.g. Kiparsky 1983, 1988, Dresher 1993, Kaisse 1993, Zec 1993), so I will discuss syn-

(2006). The point of these algorithms is to converge on the correct ranking of constraints: that is, the ranking which generates output most closely reflecting the input given to the system. These learning procedures are flexible, such that phonological change later in life is dealt with using the same mechanisms as those which bring the infant language learner from the initial state to a first language grammar. However, because these algorithms only provide procedures for reranking constraints, they are unable to alter underlying representations; at best, they can change the surface realization of an existing contrast. This may be illustrated by comparing the predicted learning experience of 1P and 2P speakers faced with the opposite low back vowel system. The 2P speaker who exhibits a surface contrast between *cot* and *caught* is assumed to have different underlying forms for these words, as well as a constraint ranking such that the Markedness (M) constraint which militates against the surface appearance of whichever phoneme is the marked member of the pair is dominated by the Faithfulness (F) constraint which enforces surface realization of the contrast. To neutralize this contrast on the surface, the 2P speaker need only rerank M and F so that M now dominates; any OT learning algorithm could easily account for this type of intraspeaker change. The situation is different for the 1P speaker faced with 2P input, however. Assuming that the IP speaker has previously internalized ambient 1P forms via Lexicon Optimization (Prince and Smolensky 1993), his underlying forms of *cot* and *caught* will be [kɑt] and [kɑt] (or [kɔt] and [kɔt]), while the relevant constraint ranking will be M $\gg$ F, reflecting both usual assumptions about the initial state (Smolensky 1996) and the lack of evidence for an opposite ranking in the ambient dialect. For this speaker, the reranking of F over M will be a vacuous change, as there is no underlying contrast to which F can enforce faithfulness. For the reranking to have an effect, the speaker still must somehow learn, for each lexical item, which low back vowel it takes.

[10]For instance, Kenstowicz 1994, p. 22 states: "We conceive of sound change as an alternation of the plus/minus specifications of the entries in the feature matrix"; it is unclear whether this alternation occurs across generations or within a particular speaker (or both).

[11]I am not assuming that intraspeaker linguistic change and intergenerational change are completely analogous. However, I think the extrapolation from community change to change over the lifespan is a valid one: whether we are talking about Speaker A's system at time T vs. A's state at T+n, or Speaker A's system vs. that of Speaker A's child B, the varieties under discussion must all be a) possible human languages and b) related in a principled way. The main drawback to this approach is that it is likely to be not restrictive enough: while a speaker who has already acquired a language will be constrained by both this language and Universal Grammar (UG) when making any changes in response to new input, the first language learner (a.k.a. the putative source of language change) is only constrained by UG.

chronic changes in terms of the levels assumed by this theory. Given Lexical Phonology (henceforth LP), there are three potential loci of phonological change in an individual's grammar: underlying (featural) lexical representations, lexical rules, and postlexical rules.

Speakers who do not have a low back vowel contrast (1P speakers) are assumed to store identical featural representations for the two lexical items *cot* /kɑt/ and *caught* /kɑt/. In order for "complete" unmerging - in the sense of replication of a 2P speaker's low back vowel output - to occur, every low back vowel in the 1P speaker's lexicon must be altered to include an additional feature that will enable later stages in the derivation (ultimately, the phonetic component) to realize the contrast. Such comprehensive acquisition of the contrast as realized in a 2P dialect seems unlikely, as Labov (1994) and others have pointed out; the would-be 2P individual may simply not be exposed to tokens of every low back vowel word in the new dialect. The unlikelihood of complete unmerging in this sense has been put forth as an argument for why mergers are necessarily irreversible Labov (1994). However, this is a straw man; clearly there is a (logically-possible) middle ground between learning a new sound for all relevant words and learning the sound for none of those words. If features can be added to underlying representations, then we might expect that these additions would occur on a word-by-word basis, with perhaps highly frequent and/or highly salient words acquiring a value for the new feature first.[12] While this change would occur in a lexically-gradual manner (in what may be termed a "split-by-transfer", in contrast to the phenomenon of

[12]This is not a prediction that falls out from the generative theory of representation directly, but arises as a sort of performance factor affecting change. Just as speakers may be prevented from understanding a sentence with multiple center-embedded clauses due to constraints on memory or attention, speakers will be hard-pressed to learn new, lexically idiosyncratic features for words they haven't heard yet.

merger-by-transfer (Trudgill and Foxcroft 1978)), the results of it ought to be phonetically-abrupt: words may vary in terms of when they receive their new feature value, but because the words will be receiving *one of two* values for that new feature, they can ultimately be phonetically spelled-out in one of two ways.

If underlying representations are indeed immutable, it may be possible for the 1P speaker to accommodate to the 2P dialect at a later stage in the phonological derivation. In LP, this can occur through two types of rules: lexical rules, may single out specific lexical items but must also be structure-preserving,[13] and postlexical rules, which are not bound by structure-preservation, but cannot admit lexical exceptions. These constraints on how rules operate ultimately restrict the LP system to a single type of rule-based accommodation. Lexical rules can handle changes which apply to specific words; upon exposure to 2P *caught* and *cot*, the IP speaker could posit a lexical rule that changes his representation of *caught* to /kɔt/. However, this split-by-transfer is not possible in LP, given the requirement that lexical rules be structure-preserving. In this case, because the 1P speaker's phoneme inventory does not already contain an /ɔ/, the lexical rule cannot change underlying /ɑ/s into /ɔ/. New segments may be introduced at the postlexical level, as these rules may introduce new segments which are not present in the lexicon; however, at this point no lexical exceptions can be made. It is possible, though, to formulate rules which change /ɑ/ into [ɔ] in specific phonetic environments. This will not result in a contrast, but it can result in new allophones. For instance, if the 1P speaker observes that the /ɑ/ in words like *mall, Paul* and *call* is realized as [ɔ] in the 2P dialect, then she may posit a postlexical rule which introduces an [ɔ] allophone in prelateral contexts. This account predicts use of [ɔ] in words

[13]Because they apply within the lexicon, lexical rules cannot introduce segments which are not part of the phonemic inventory; this constraint is known as *structure-preservation*.

which should not, on the basis of the 2P input, take a new segment. For example, while most 2P words containing a prelateral low back vowel have /ɔ/ as that vowel, there are a few exceptions, such as doll /dɑl/ and golf /gɑlf/. If the 1P speaker, having heard forms such as [mɔl]/[pɔl]/[kɔl], posits a postlexical rule that creates a rounded, higher low back allophone before tautosyllabic /l/, then we would expect this rule to be applied to *doll* and *golf* as well, yielding [dɔl] and [gɔlf].[14] Crucially, words which form minimal pairs should not change, since these are by definition cases in which the appearance of one sound over the other cannot be predicted by rule.

3.4.2 Contrast acquisition in usage-based phonolology

Usage-based phonology has a much more clearly defined account of intraspeaker change; this is, of course, because dynamic phonological representations are at the core of this type of theory. In the generative model, the underlying representation of a word is abstract and mostly fixed, with later rules left to do most of the heavy lifting in terms of variation and change. However, in a usage-based model, the word-level representation is the primary locus of change: new tokens of words cause shifts in their associated exemplar clouds, and changes at the level of phonology (which comprises generalizations over these word forms) merely follow from these changes in representations.

Accommodation to a new contrast is thus predicted to occur in a very different manner from that described in the previous section. In this case, the 1P speaker

[14]This behavior might be termed *hypercorrection*, if such speakers are orienting to what they see as a prestige variety; however, this motivation is unlikely, given the stigmatization of varieties spoken in and around New York City (Labov 1966, Preston 1999). As DeCamp (1972) points out, "hypercorrection is always a function of rule generalization, though the converse does not hold" (p. 1).

starts out with two lexical items, *cot* and *caught*, each of which is associated with a cloud of exemplars. Unlike those of the 2P speaker, these clouds are largely coterminous in the phonetic space. Crucially, however, they are still considered to be two distinct clouds, because tokens must be associated with one semantic label or another.[15] If the 1P speaker is exposed to a dialect in which these words are pronounced [kɔt] and [kɑt], tokens of these words will continue to be correctly stored in the relevant cloud, with the result that the clouds will ultimately start to diverge. As in the generative account, this should occur (at least initially) on a word-by-word basis, such that words to which the speaker has the greatest amount of exposure will change first. However, the usage-based theory predicts that this change should also be phonetically gradual: words are not dropped into one category or the other, but instead are expected to shift gradually in the phonetic space, reflecting the ongoing incorporation of gradiently variable heard tokens into representational clouds laden with older remembered exemplars.

3.5 Phonological generalizations

While representations are a key part of any phonological theory, the ultimate raison d'être of the phonological component in a theory of language is the existence of phonological generalizations. If there were no evidence that speakers have tacit knowledge of "rules", then there would be no need for a phonological component: the syntactic module simply could pluck fully-formed words from the lexicon and concatenate them as tree structure dictated. Of course, there is ample evidence for the existence of generalizations in phonology, not just as static descriptions

[15] It is not the case that 1P speakers are regularly stymied by utterances such as *I [kat] the ball yesterday* or *She slept on the [kat]*; in most cases, the correct semantic label is assigned.

of what exists in the lexicon, but as productive elements: children experience a period of overgeneralization as they acquire language, indicating the existence of generalizations which have been overapplied, and morphologically complex nonce words elicited from speakers of any age are phonologically law-abiding (Berko 1958).

As noted above, the two types of theories under consideration differ with respect to how these productive generalizations are formalized within the theory. In generative frameworks, rules are essentially functions which take the underlying representation as input and derive a representation which is closer to the surface form. They are separate from underlying forms, and affect entire classes of lexical items containing the relevant underlying or intermediate form. Some have proposed mechanisms for allowing lexical exceptions to rules, as noted in footnote 3; for the most part, however, rules apply across the board.

In usage-based frameworks, however, productive generalizations have a very different status. They are not separate from underlying representations, but emerge from those representations: the targets for entirely new instances of a category V in the context C are calculated at each production by averaging over the stored exemplars of V/C. However, tokens of specific words which have already been heard are derived from the existing examples of that word. While generalizations of phoneme or allophone level categories may be available in the absence of a word-level category label to activate, the word-level is primary.

These different treatments of generalizations mean that the generative and usage-based theories make very different predictions about how speakers who natively exhibit a phonological generalization such as Canadian Raising might go about accommodating to a dialect which is not characterized by this generalization. How can speakers "lose" a rule in each of these frameworks?

3.5.1 Altering a rule in generative phonology

In a generative account, Canadian Raising is formalized as in (1):

(1) /aw/ → [ʌw] / _ [-VOICE]

This rule changes underlying /aw/ into [ʌw] when it occurs before voiceless sounds; instances of /aw/ which do not meet this structural description are sent forward through the derivation as [aw]. The difference between [ʌw] and [aw] is ultimately spelled-out as a difference in vowel height, with [ʌw] realized higher in the vowel space than [aw]. Speakers of NYAE, who do not exhibit Canadian Raising, are assumed to have no such rule; for these speakers, all underlying /aw/ pass through to phonetic implementation as [aw], with the result that no phonologically-based distinction is made between allophones of these segment.

Given a system that contains the rule in (1), accommodation to a variety such a NYAE which does not exhibit raising may happen in one of three ways. One possibility is that the rule may simple cease to apply, resulting in a system which is qualitatively the same as that of NYAE; the result of this change would be phonetically and lexically abrupt, such that all underlying /aw/s would receive the same nonraised phonetic realization. Alternatively, the rule may begin to variably apply: prevoiceless tokens of (aw) would sometimes be realized as raised, sometimes unraised, but importantly, this variability would affect all lexical items as such to the same degree. Finally, the raising rule may be maintained, but the phonetic realization of the resulting [ʌw] might be adjusted to be somewhat less raised than it would be in CE.[16]

While the phonetic results of these possibilities differ, the lexical impact is predicted to be the same: because each change involves an adjustment to one rule which affects all pre-voiceless (aw) lexical items to the same extent, no word-specific effects should occur.

3.5.2 Altering a rule in usage-based phonology

In a usage-based account, the Canadian Raising generalization emerges from all the representations which instantiate it. According to this view, a speaker of CE does not have a single rule of Canadian Raising, but a multitude of exemplars of /aw/-containg words. These exemplars forms a bimodal distribution in the parametric phonetic space, with prevoiceless and nonprevoiceless tokens of (aw) forming distinct dense clouds. Given this situation, the process of "losing" a rule is much more gradual: relevant lexical items must gradually shift in the phonetic space as new exemplars of them are encountered. Such shift will occur on a word-by-word basis, with more frequent lexical items acquiring new exemplars at a faster rate than rarer items.

The process is essentially the reverse of that which occurs in acquiring a new contrast in usage-based theory. Rather than gradually splitting an existing single distribution into two, the elimination of raising in this case involves gradually collapsing two existing distributions into one. In both cases, lexical frequency

[16]Some combination of these three basic changes is also possible. For example, the Canadian Raising rule might be made variable, with later phonetic implementation rules realizing [ʌw] with a somewhat lower nuclear quality.

effects are predicted: more frequent items will be the first to shift towards new dialect realizations.

3.6 Accommodation to both features

The generative and usage-based accounts described here differ in several ways with regards to the types of changes that are predicted to occur when Canadian speakers are exposed to NYaE. As noted in previous sections, each type of theory makes specific predictions about how a *cot/caught* contrast may be acquired, and how a Canadian Raising generalization may be lost or changed. However, each type of theory also has implications for how these two features might be changed with respect to one another.

In generative theory, if speakers are to accommodate to the NYaE low back vowel contrast, this will occur on a phonetically-abrupt, word-by-word basis. This should be a slow process, as it will involve acquiring sufficient input for each of the several hundred lexical items containing the single low back of CE. Accommodating to the low prevoiceless (aw) of NYaE is, however, is a comparatively simple change, involving the alteration of merely one element in the phonology, the raising rule. This asymmetry implies that speakers exposed to NYaE are more likely to alter the raising rule than they are to acquire the low back vowel contrast.

In usage-based theory, both contrast acquisition and raising-loss are predicted to occur in a phonetically and lexically gradual manner. This is because both occur via the same mechanism: gradual shift of individual word clouds as a result of the gradual addition of new dialect exemplars. Because these changes are implemented in the same way, there is no predicted asymmetry in terms of ease of acquisition; speakers are not expected to have any more trouble acquiring a contrast than they do losing a rule. In this account, we would predict that accommodation to these two features would be present to the same extent within speakers: those who have successfully changed their (aw)-raising rule are also likely to have accommodated to the low back vowel contrast.

In chapters to follow, I test these predictions by analyzing the production of each feature in turn (low back vowel systems in chapter 4, Canadian Raising in chapter 5), then examining whether any relationship exists between the two features within speakers (chapter 6).

CHAPTER 4

LOW BACK VOWELS

Contrast is a central concept in any phonological framework. In the last chapter I argued that the acquisition of a new contrast in the course of second dialect acquisition potentially provides an important test of competing phonological theories. This chapter presents the results of such a test and discusses its theoretical implications: (how) do native speakers of Canadian English accommodate to the low back vowel distinction of New York area English, and what theory best accounts for these facts?

Chapter 2 contained a brief review of the current status of the low back vowels in North American English, with special attention to the varieties spoken in Canada and the New York City region. In this chapter, I outline the specific linguistic questions raised by the potential acquisition of a low back vowel contrast, and detail the methods of phonetic and statistical analysis used to address them. I then present the findings for each individual speaker before attempting to make generalizations across speakers, and end the chapter with a discussion of these findings in light of the phonological predictions made in Chapter 3.

4.0 Low back vowels: Three perspectives

Before embarking on an analysis, it is important to set out exactly what is meant by the phrase "low back vowels". Three conceptually distinct though intertwined perspectives can be teased apart:

Low back vowels as phonetic patterns. This is the least abstract[1] level from which to view the sounds of language. Articulatorily speaking, "low back vowel" refers to any vowel sound which is produced with the tongue in a low and backed position in the oral cavity; acoustically speaking, this label could be applied to any signal with the high first formant and low second formant (and associated placement in the vowel space) that result from that articulatory configuration. Both of these descriptions accurately delimit the set of physical tokens subjected to analysis in this chapter, and both are ultimately crucial to understanding accommodation, which involves hearing and producing new forms in addition to mentally encoding them.

Low back vowels as phonological entities. It is also possible to think of these vowels more abstractly, as entities in the mental lexicon and grammar. Depending on one's theoretical perspective on the nature of these objects, referring here to "low back" vowels at all might be considered a category error: these adjectives may describe the physical realizations corresponding to abstract phonological entities, but are arguably inappropriate for mental representations which consist not of actual tongue gestures or sound waves, but patterns of neurons. That said, it is almost impossible to define these vowels in a way that removes all phonetic

[1] Though still abstract: generalizations about the types of articulatory or acoustic patterns characterizing the objects identified as low back vowels do, after all, involve a certain amount of abstraction away from particular physical instances.

substance: even the formal features used to pick them out from the rest of the phonemic system in generative phonology ([LOW],[BACK]) make reference to their eventual phonetic spell-out. For the time being, it is enough to stipulate that there exist one or more mental representations which map onto (but are ontologically separate from) the phonetically defined low back vowel space. These representations could be given very abstract labels such as X and Y, but the conventional labels /ɑ/ and /ɔ/ have been adopted in the discussion thus far. In generative phonology, these are the sole underlying representations for these vowels; in usage-based theories, these representations correspond to clouds of stored exemplars, and may be formalized as linguistic category labels that are indexed to these exemplars.

Low back vowels as word classes. Like all segments, low back vowels are not encountered in isolation in normal speech; rather, they nearly always occur in combination with other phonemes.[2] The particular set of words that contains a particular (phonological) vowel is ultimately arbitrary. That is, one cannot predict based on semantic content whether a word will contain, for example, an /i/; this is the insight captured by Saussure's notion of the arbitrariness of the sign (de Saussure 1916).

However, these sets tend to hang together over time: for example, {English Words Containing /i/ 50 Years Ago} and {English Words Containing /i/ Today} are largely, if not completely, coextensive. The most well-known classification of these cohesive word groups in English is the Standard Lexical Sets of Wells (1982), which lays out 24 word classes based on the quality of the vowel in their stressed

[2]Possible exceptions involving the low back vowels might be the affective vocalizations *Aw!* [ɔ] and *Ah!* [ɑ], but these are probably better analyzed as something like free morphemes which happen to consist only of 1 phoneme, similar to the bound inflectional morphemes -*z* and -*d*, which also coincidentally consist only of 1 phoneme.

syllable[3] as realized in two reference accents, Received Pronunciation and General American.

Of course, changes may occur to these sets: words which at some previous point contained the same vowel sound may split into two or more groups, sets may merge, or individual lexical items may hop to a new word class. However, the general historical integrity of these lexical sets provides a useful way of talking about vowels at a level of abstraction somewhere between the phonetic and phonemic levels described above. By referring, for example, to the realizations of the vowels in NORTH versus FORCE, the linguist can generalize beyond particular tokens of the words in these sets, while remaining agnostic about their status relative to one another in the mind of any actual speakers. In addition, they provide a historically grounded starting point for an analysis which aims to get at more strictly phonological questions, by pointing out which words might be expected to pattern together and which might not.

There are several lexical sets which are potentially relevant to a study of low back vowels, insofar as the words included within these sets have a vowel whose quality places it within the phonetic low back space. These sets are summarized in Table 4.1.

NORTH and START will not be included in the current study. These two sets are potentially interesting in a contact study of Canadian English and American English, as a particular set of START words - those with the vowel appearing before an ambisyllabic /r/ - have taken on the NORTH vowel in Canadian English. However, the more general merger of /ɔ/ and /ɑ/ does not affect these sets, so

[3]There are also 3 lexical sets dealing with unstressed vowels, HAPPY, LETTER, and COMMA. While not part of the standard 24, variation in the unstressed vowels represented by these sets has been investigated in recent studies (e.g. Harrington 2006, Watt et al. 2010).

Set	Received Pronunciation	General American
THOUGHT	ɔː	ɔ
CLOTH	ɒ	ɔ
LOT	ɒ	ɑ
PALM	ɑː	ɑ
NORTH	ɔː	ɔr
START	ɑː	ɑr
CHOICE	ɔɪ	ɔɪ

Table 4.1: Lexical sets containing low back vowels, and their typical realizations in Received Pronunciation and General American English (Wells 1982)

they will be left for future study. In addition, the CHOICE set will be ignored in the remainder of this dissertation. While the quality of /ɔɪ/'s nucleus may be phonetically described as low back, the diphthong as a whole is not confined to this space. More importantly, it is relatively uninteresting from a phonological standpoint within the context of the current study, as CHOICE is clearly distinct from all the other low back sets in CE as well as NYAE, and the mapping of CHOICE between these varieties is straightforwardly one-to-one.

This study will focus on realization of the THOUGHT, CLOTH, LOT, and PALM sets, as these are the sets which participate in the low back vowel merger. As indicated in Table 4.1, these four sets fall into a two-way contrast in many dialects of American English: THOUGHT and CLOTH are produced with a higher, backer vowel, which contrasts with the lower, fronter vowel of LOT and PALM. Through the rest of this chapter, I will follow the conventions of Labov et al. (2006) (and much of the literature on variation and change in North American English) and use the symbols (oh) and (o) to refer to the potentially contrastive supersets THOUGHT/CLOTH and LOT/PALM, respectively.

In this chapter I will talk about low back vowels from all three of these perspectives. I am primarily concerned with the phonological perspective: do my speakers have 1 or 2 abstract low back vowel categories? However, drawing these phonological conclusions will require examining the distribution of tokens in phonetic space and to what extent these distributions coincide with historical word class.

4.1 The research questions

The first question addressed in this chapter is whether there any evidence that the Canadian speakers in this study have acquired a NYaE-like contrast between (o) and (oh). That is, can it be shown that these speakers produce a measurable phonetic distinction between words in these classes that is not predictable by phonological context, but instead predictable only by the lexical distribution of these sounds in the ambient dialect? The answer to this first question, of course, may depend heavily on social factors: a speaker may show no evidence of having acquired this feature, for instance, due to a general unwillingness to accommodate or lack of sufficient interaction with speakers of the new dialect. The effect of such extralinguistic factors will be examined in Chapter 7.

Also relevant to the linguistic questions raised in Chapter 3 is the manner in which this acquisition takes place. If speakers show evidence of a contrast as described above, what seems to be the mechanism of split underlying this change? Are there frequency or other word-specific effects on vowel quality in the expected directions? Do the resulting two categories occupy two distinct phonetic territories, or is there evidence of phonetic gradience?

To address these question, I will analyze three types of production data: minimal pairs, word lists, and conversation.

4.2 Minimal pairs

As described in Chapter 2, speakers were asked to read and evaluate a series of minimal and rhyming pairs, many of which probed the (o)/(oh) distinction: speakers read aloud a pair such as *caught/cot* or *pot/bought* and then stated whether they thought the pair sounded the same (for minimal pairs) or rhymed (for rhyming pairs). This task thus yielded production data as well as perception data.

4.2.1 Acoustic analysis

Measurements of F1 and F2 were taken for each low back vowel token at the F1 maximum, which is the point representing the lowest point of the vowel. Each word pair was then checked to ensure that the measurement points for the two words in the pair represented reasonable points of comparison.The duration of each vowel was also measured on a pairwise basis, to ensure that the same landmarks were used to establish the beginning and end of the vowel.

4.2.2 Statistical analysis

To assess whether each speaker produces a distinction between (o) and (oh) in the minimal pair style, **paired t tests** were used to compare F1, F2, and duration across the two word classes. Paired t tests may be used in studies which compare the performance of individual subjects on some measure before and after an intervention, or in two different research conditions. An example of this might be a study of reaction time before and after drinking coffee. In such a study, some small effect of caffeine consumption on reaction time (presumably, in the direction of decreasing it) might be predicted. However, a small difference between experimental conditions might be dwarfed by the differences between subjects, who will vary in their reaction times regardless of caffeine consumption: some speakers will be relatively quick in both conditions, while others will be slow. Because of the large interspeaker variability, if we simply use an unpaired Student's t to test for a difference in mean reaction times between all subjects before coffee drinking and all subjects after coffee drinking, it is likely that no significant effect will be found. The paired t test, in contrast, controls for this cross-subject variation by instead pairing the observations associated with each subject and assessing whether the mean of all the differences between pairs is significantly different from zero. If there is no effect of condition on the measured value - that is, if the difference between conditions is sometimes positive and sometimes negative, to more or less the same extent - then these differences will average out to about 0. If there *is* an effect of condition on the measured value, then this mean difference will be significantly different from 0.

Paired t tests are thus ideal for cases in which the *between*-group variation is small compared to the variation *within* those groups. This, of course, is exactly the situation faced in determining whether the CE speakers in this study are producing a (o)/(oh) distinction in their minimal pairs: the difference between the two word classes is likely to be slight, while the differences across pairs due to varying phonological contexts is likely to be great. Using the more powerful paired t test increases the likelihood that any difference between the (o) and (oh) words in this list will be detected.

4.2.3 Minimal pair results

4.2.3.1 Perception.

All speakers uniformly reported that the THOUGHT/LOT pairs sounded "the same" after producing them. In the parlance of the sociolinguistic literature on merger, they all exhibit a "merger in perception" with respect to these two word classes.

However, there was variation among responses for the THOUGHT/PALM pairs elicited (*law/la* and *paw/pa*), relating directly to the way in which each speaker pronounced *pa* and *la*. Several speakers produced *pa* and *la* with a fronted, more [a]-like vowel, indicating that these words occupied a different word class.[4] There was similar variation in the LOT/PALM pairs: while all speakers agreed that *bother* and *father* rhyme, some speakers detected a difference in *logger/lager*, which was dependent on whether the speaker produced *lager* with a fronter, [a]-like vowel. Unsurprisingly, all speakers said that *daughter/water* rhymed, as this is a rhyming pair in both CE and NYaE.

4.2.3.2 Production.

Paired t tests were used to compare F1, F2, and duration values between (o) and (oh) words across the minimal and rhyming pairs. The THOUGHT/PALM and LOT/PALM pairs were excluded from analysis due to the variable treatment of particular PALM class words on the list, as discussed in the previous section. The pair *daughter/water* was also excluded, as it includes two words from the (oh) class.

[4] Possibly the "foreign a" class described in Boberg (2009)

The results of paired t tests comparing F1, F2, and duration for (o) and (oh) in minimal pair style are presented in Tables 4.2, 4.3, and 4.4.[5]

No significant difference was found for any measure[6] between the two vowels in this style, with one exception: JC's (oh) and (o) differ significantly in F2 ($t(9) = -2.6664$, $p=0.03$). This single significant result may very well be a chance occurrence: given the number of tests run across speakers, two or three such significant results might be expected.[7] However, it may also be grounded in the particular linguistic history of this speaker, whose father was born in Brooklyn.

Aside from JC, however, 16 of the 17 speakers show a merger in production consistent with their merger in perception. In this style, at least, they do not seem to be showing much accommodation towards the NYaE contrast.

[5]Means, standard deviations, and mean of differences are all rounded to the nearest Hertz or millisecond. P-values are rounded to two decimal places. For each of speakers BK, CW, SS, and VJ, one of the 9 (oh)/(o) minimal pairs had to be excluded, because the formant measurements for at least one token of the pair were not reliable; this accounts for the variation in degrees of freedom.

[6]One striking pattern which does emerge from this data is the overall trend in duration patterns exhibited across speakers. While no individual speaker has a significant duration difference between (o) and (oh) words, every speaker's (oh) class is slightly longer than her (o) class, an effect which is consistent with phonetic differences between these vowels in NYaE.

[7]Three tests (of F1, F2, and duration) for each of 17 speakers = 51 separate tests. Using the .05 criterion, normally distributed data would yield approximately 2.5 "significant" results due to chance alone.

	o F1 (Hz)		oh F1 (Hz)				
Speaker	mean	SD	mean	SD	mean difference	t(df)	p
BK	746	98	727	125	20	t(8) = -0.5432	0.60
BW	624	25	626	30	-2	t(9) = 0.2839	0.78
CW	805	27	793	58	12	t(8) = -0.7821	0.46
DB	685	46	696	67	-12	t(9) = 0.8040	0.44
ES	614	58	595	79	19	t(9) = -0.7242	0.49
EW	612	31	608	58	4	t(9) = -0.2481	0.81
GH	713	68	708	75	6	t(9) = -0.4123	0.69
JC	652	61	645	58	7	t(9) = -0.8615	0.41
JF	684	47	693	47	-9	t(9) = 1.0446	0.32
LC	782	66	793	90	-11	t(9) = 0.6808	0.51
LG	722	80	746	81	-25	t(9) = 1.0737	0.31
LW	758	42	764	91	-5	t(9) = 0.1596	0.88
NW	745	114	744	110	1	t(9) = -0.0403	0.97
PW	669	57	654	46	15	t(9) = -1.1469	0.28
SS	670	72	660	77	10	t(8) = -0.6105	0.56
TM	766	55	737	63	29	t(9) = -1.3920	0.20
VJ	661	74	656	145	4	t(8) = -0.1125	0.91

Table 4.2: (o)/(oh) minimal pair production results: F1. Mean and standard deviations for each word class as produced in Minimal Pair readings, with paired t test results.

	o F2 (Hz)		oh F2 (Hz)					
Speaker	mean	SD	mean	SD	mean difference	t(df)		p
BK	1127	205	1124	184	3	t(8) =	-0.0345	0.97
BW	1010	63	1003	60	7	t(9) =	-0.5399	0.60
CW	1149	49	1116	61	33	t(8) =	-1.3901	0.20
DB	1019	76	1032	93	-13	t(9) =	0.7717	0.46
ES	1055	132	1030	88	25	t(9) =	-0.8064	0.44
EW	924	60	929	83	-5	t(9) =	0.2381	0.82
GH	1095	95	1086	110	9	t(9) =	-0.6101	0.56
JC	984	79	953	82	31	t(9) =	-2.6664	0.03
JF	1022	64	1040	52	-17	t(9) =	1.0293	0.33
LC	1014	58	1078	157	-64	t(9) =	1.2168	0.25
LG	1017	51	1009	100	8	t(9) =	-0.2761	0.79
LW	1142	97	1121	107	21	t(9) =	-0.5767	0.58
NW	1180	77	1181	90	-1	t(9) =	0.0294	0.98
PW	1048	107	1009	61	39	t(9) =	-2.0037	0.08
SS	1101	49	1087	72	14	t(8) =	-0.9279	0.38
TM	1298	169	1223	154	75	t(9) =	-1.8134	0.10
VJ	1098	78	1117	65	-19	t(8) =	0.6641	0.53

Table 4.3: (o)/(oh) minimal pair production results: F2. Mean and standard deviations for each word class as produced in Minimal Pair readings, with paired t test results.

	o duration (ms)		oh duration (ms)					
Speaker	mean	SD	mean	SD	mean difference	t(df)		p
BK	164	42	169	64	-5	t(8) =	0.3636	0.73
BW	212	50	214	47	-2	t(9) =	0.1632	0.87
CW	235	59	263	87	-27	t(8) =	1.0067	0.34
DB	263	70	270	58	-7	t(9) =	0.5324	0.61
ES	234	74	236	84	-2	t(9) =	0.2073	0.84
EW	214	46	215	61	-1	t(9) =	0.0658	0.95
GH	211	57	218	61	-6	t(9) =	0.6543	0.53
JC	210	73	216	90	-5	t(9) =	0.4566	0.66
JF	187	75	190	71	-2	t(9) =	0.2244	0.83
LC	244	70	257	89	-14	t(9) =	1.3847	0.20
LG	198	64	202	60	-4	t(9) =	0.4251	0.68
LW	248	65	264	84	-16	t(9) =	1.0386	0.33
NW	235	79	236	75	-1	t(9) =	0.0613	0.95
PW	233	78	241	87	-8	t(9) =	0.4961	0.63
SS	264	98	286	113	-22	t(8) =	2.1928	0.06
TM	210	55	235	82	-25	t(9) =	1.5736	0.15
VJ	209	71	237	138	-28	t(8) =	0.6088	0.56

Table 4.4: (o)/(oh) minimal pair production results: Duration. Mean and standard deviations for each word class as produced in Minimal Pair readings, with paired t test results

(oh)	(o)
bought	pot
cloth	lot
coffee	copy
pawn	don
taught	dot
thought	
water	

Table 4.5: (o)/(oh) word list 1.0: Words elicited from BK, GH, JC, SS, and VJ

4.3 Word list

The original point of the word list in this study was to elicit a few tokens of every lexical set, with the aim of establishing a citation form vowel space. The first version of the word list, administered to the first 5 speakers interviewed, contained 7 (oh) words and 5 (o) words (Table 4.5). Measurement points for these vowels were taken in the same way as for the minimal pair list tokens.

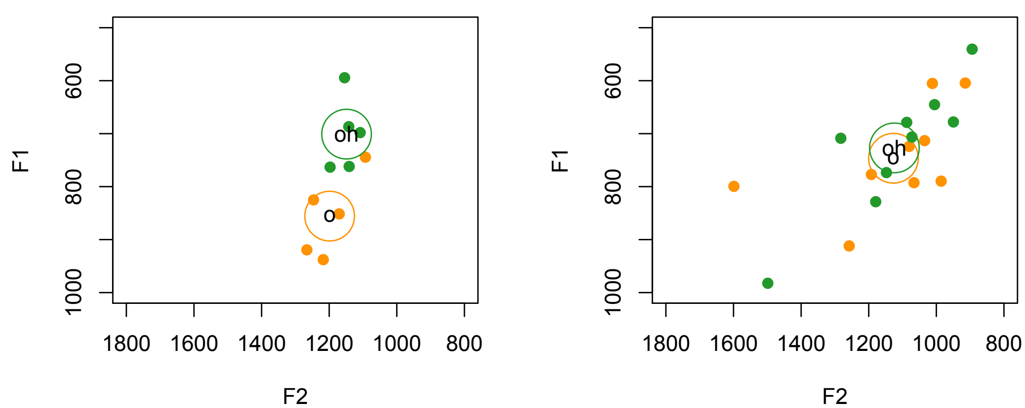
It became apparent that, for some of these speakers, there seemed to be an effect of style on low back vowel production: for speakers BK, GH, SS, and VJ, the vowels were auditorily more distinct in Word List style, and showed greater separation in the vowel space (See Figures 4.1, 4.2, 4.3, 4.4, and 4.5[8]). Though a significant difference between (o) and (oh) in either dimension could not be established given the small number of tokens for these speakers,[9] these impressionistic results indicated the need for a more deliberate investigation of the low back vowel contrast in Word List versus Minimal Pair style.

[8]Word list plots for these speaker exclude *thought* and *water*; only the 5 near-minimal pairs *bought/pot*, *cloth/lot*, *coffee/copy*, *pawn/don*, *taught/dot* were plotted, for balance.

[9]In fact, 2 tests approach significance: BK's f1, and GH's f2, which is reflected in the plots of the vowels for these speakers.

(oh)	(o)
awed	odd
bought	pot
caller	collar
caught	cot
cloth	lot
coffee	copy
dog	cog
dawn	don
naughty	knotty
pawned	pond
taught	dot
talk	sock
tall	doll
walk	wok

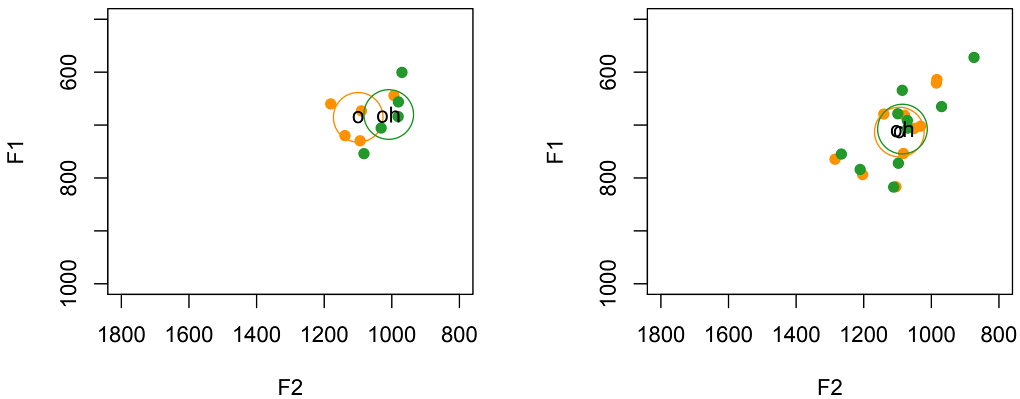
Table 4.6: (o)/(oh) word list 2.0: Words elicited from the last 12 speakers



(a) Word List

(b) Minimal Pairs

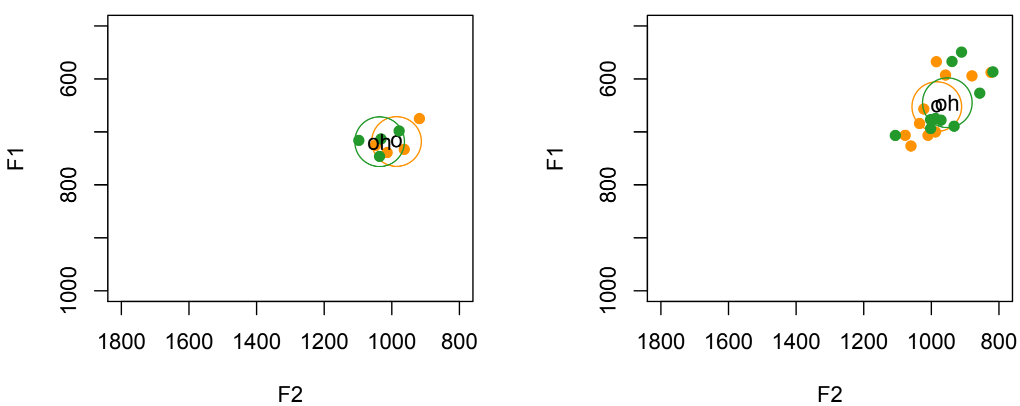
Figure 4.1: BK's (oh) and (o) in read styles



(a) Word List

(b) Minimal Pairs

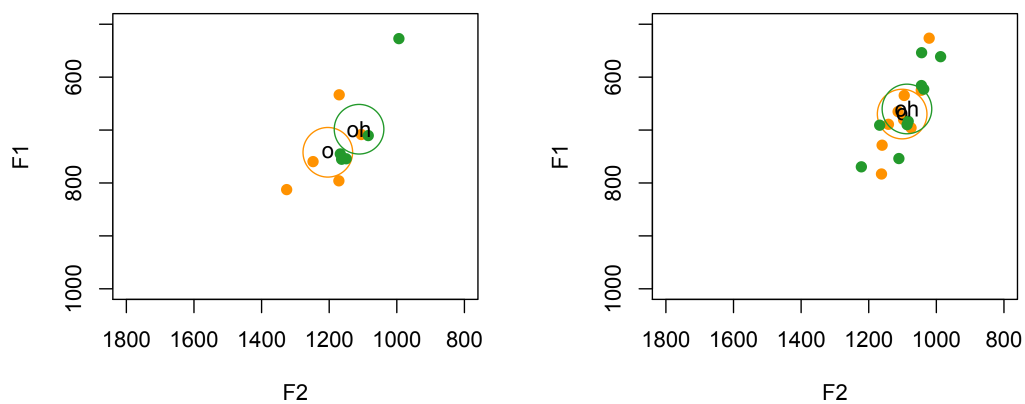
Figure 4.2: GH's (oh) and (o) in read styles



(a) Word List

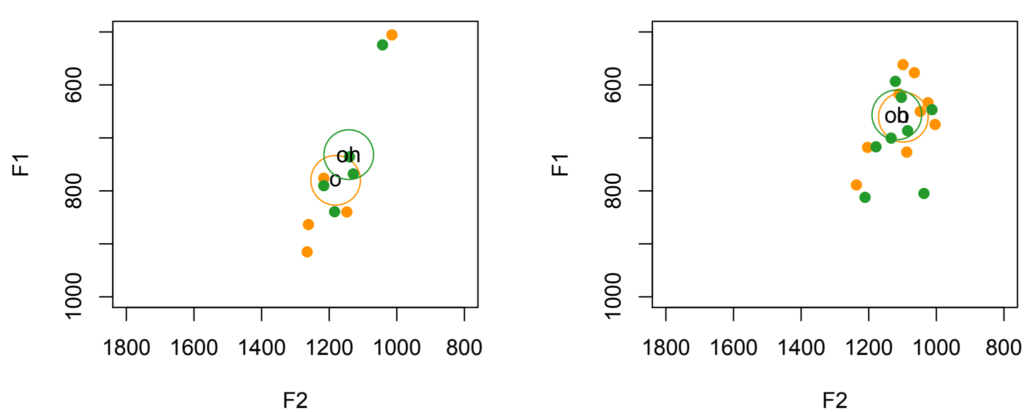
(b) Minimal Pairs

Figure 4.3: JC's (oh) and (o) in read styles



(a) Word List (b) Minimal Pairs

Figure 4.4: SS's (oh) and (o) in read styles



(a) Word List (b) Minimal Pairs

Figure 4.5: VJ's (oh) and (o) in read styles

To this end, the word list was expanded to include all of the low back vowel word pairs already included in the minimal pair list (4.6). This change enabled a statistical examination of whether a contrast was present in the word list style alone, as well as a comparison of words across styles to see whether a shift had taken place in one or both vowels.

Formant and duration measurements were taken as described above for each of the Word List 2.0 tokens, and paired t tests used to determine whether there was a significant difference between the word classes in any measure for any speaker.

4.3.1 Word list results

Several patterns of results were found among this group of twelve speakers; these results are summarized in Tables 4.7, 4.8, and 4.9.

For BW, DB, EW, and LC, no significant difference was detected in any dimension between (o) and (oh) in Word List style. There also appears to be no appreciable shift in vowel quality across read styles. For these speakers, the two word classes occupy essentially the same vowel space in both Word List and Minimal pair styles (see Figures 4.6, 4.7, 4.8, and 4.9).

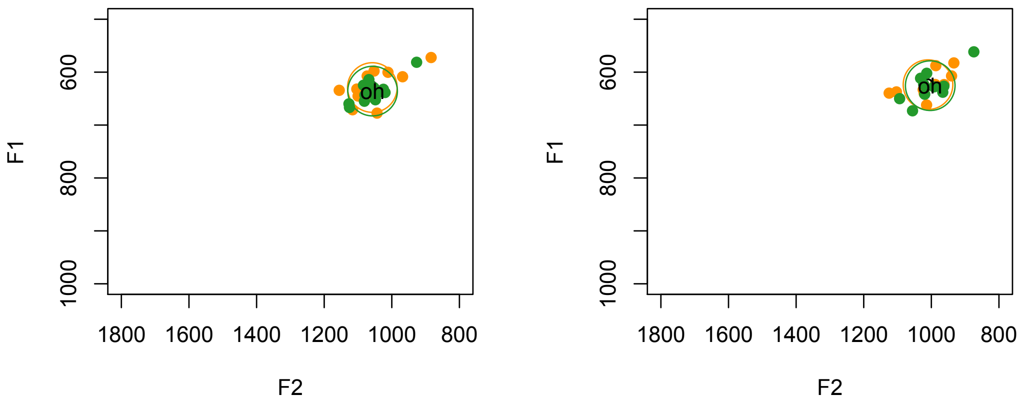
Speakers ES and LW showed no significant difference between word classes in either formant measure, though ES's (oh) is significantly longer than his (o) in Word List style. However, while these speakers do not seem to distinguish two vowels in either Minimal Pair or Word List style, there does seem to be a shift in vowel quality across these styles: for both speakers, their apparently single low back vowel is fronter and lower in Word List style than in Minimal Pair style (see Figures 4.10, and 4.11).

JF also shows no significant difference in Word List style. However, her (o) seems to be slightly (if not significantly) lower in Word list style compared to Minimal Pair style (Figure 4.12)

Speakers CW and NW both show a significant difference in F1 between (oh) and (o). In both cases, it appears that (o) is lower in Word List style than in Minimal Pairs; (oh), meanwhile, does not seem to shift (Figures 4.13 and4.14).

Finally, speakers LG, PW, and TM show significant differences in F2 between (o) and (oh) in Word List style. Again, this difference seems mainly to be due to a shift in (o) across contexts, though LG's (oh) also appear to be somewhat backer in Word List context (Figures 4.15, 4.16, and 4.17).

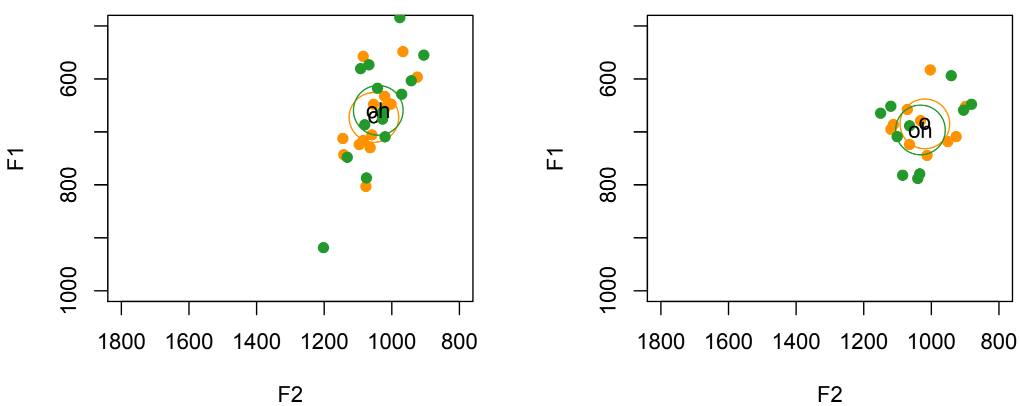
To summarize, half of the 12 speakers who read the second, fuller Word List distinguished between (o) and (oh) in this style. None of these speakers distinguished these vowels in the Minimal Pair style, so clearly some shift in one or both word classes must have taken place. Indeed, a mere visual comparison of the vowel plots for each of these styles across speakers indicates that those speakers who have shifted have done so in a consistent manner: 6 speakers seem to be producing their (o) word class in a fronter and/or lower (i.e. more (o)-like) position, while 2 speakers who did not distinguish (o) and (oh) in Word List nonetheless produce this single undifferentiated vowel in a fronter and lower position.



(a) Word List

(b) Minimal Pairs

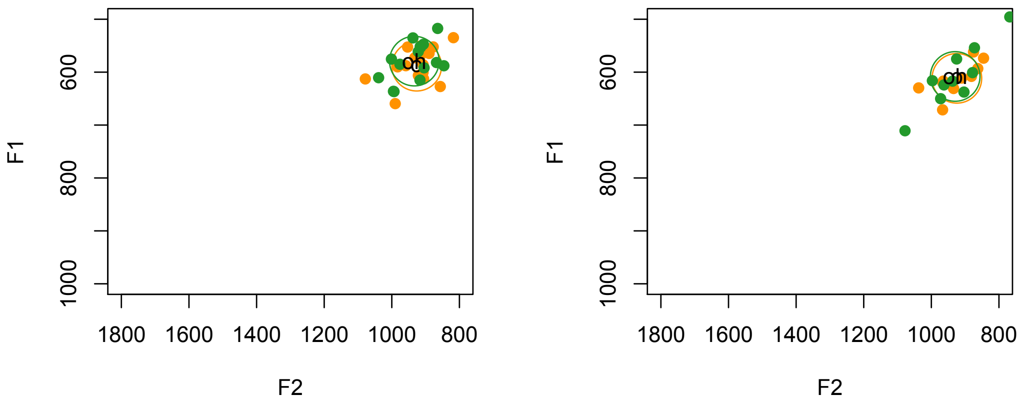
Figure 4.6: BW's (oh) and (o) in read styles



(a) Word List

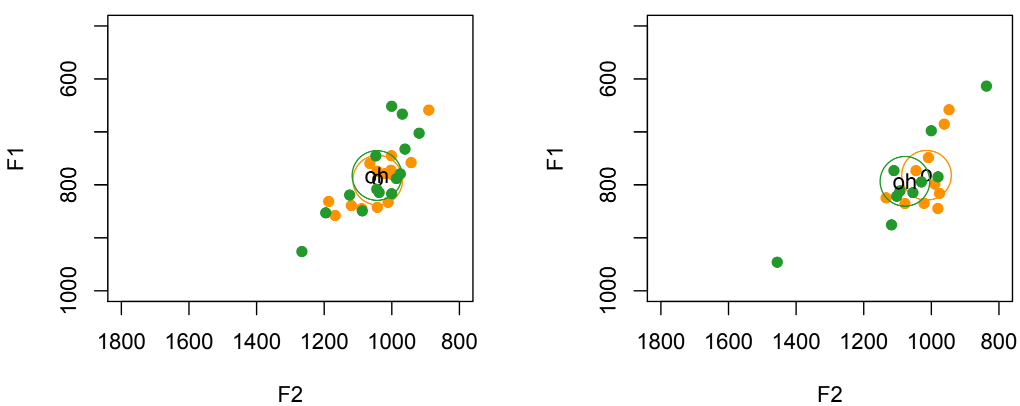
(b) Minimal Pairs

Figure 4.7: DB's (oh) and (o) in read styles



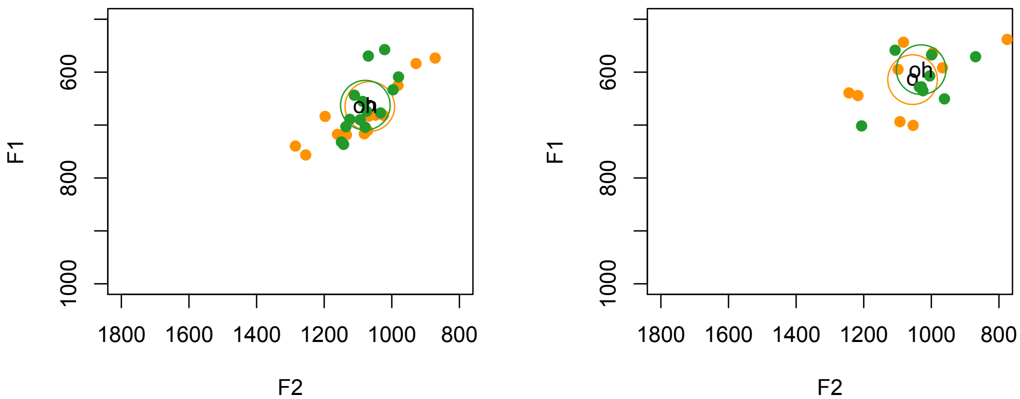
(a) Word List (b) Minimal Pairs

Figure 4.8: EW's (oh) and (o) in read styles



(a) Word List (b) Minimal Pairs

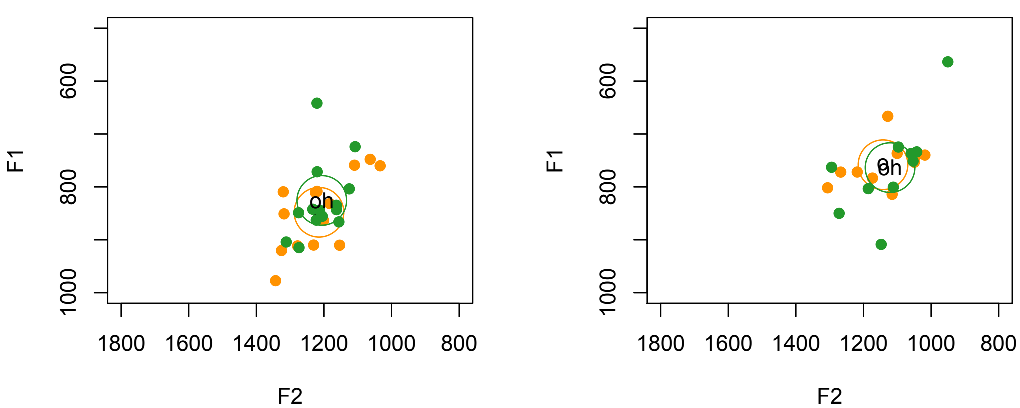
Figure 4.9: LC's (oh) and (o) in read styles



(a) Word List

(b) Minimal Pairs

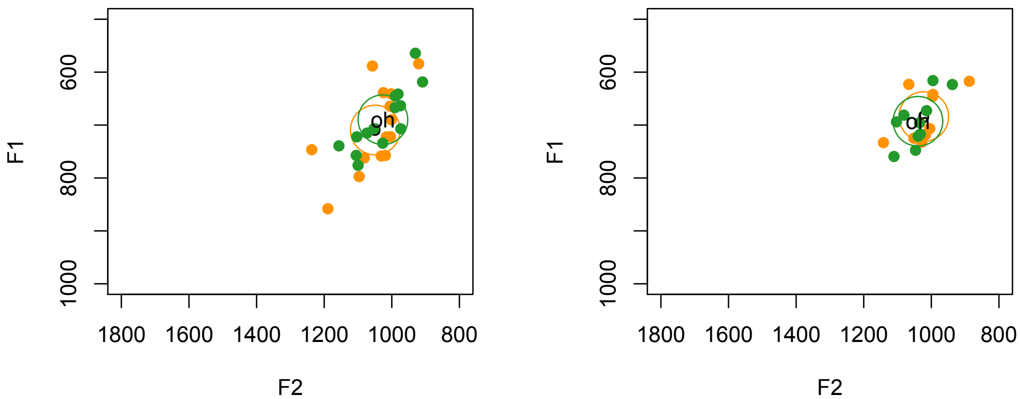
Figure 4.10: ES's (oh) and (o) in read styles



(a) Word List

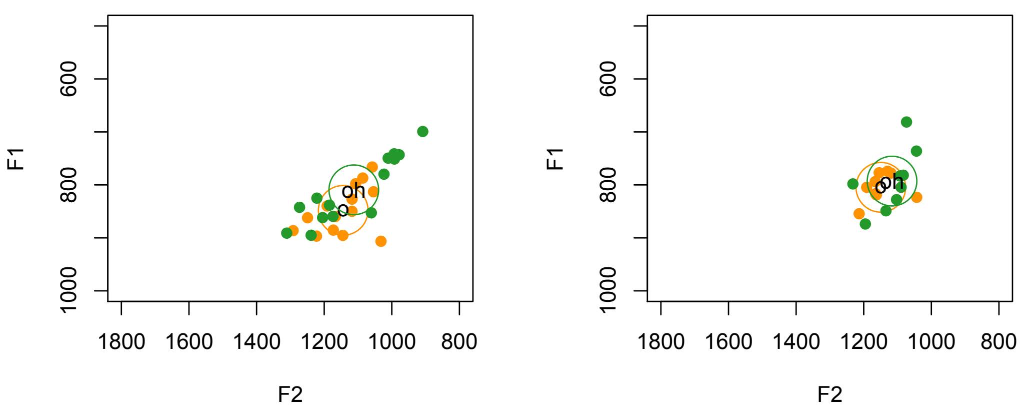
(b) Minimal Pairs

Figure 4.11: LW's (oh) and (o) in read styles



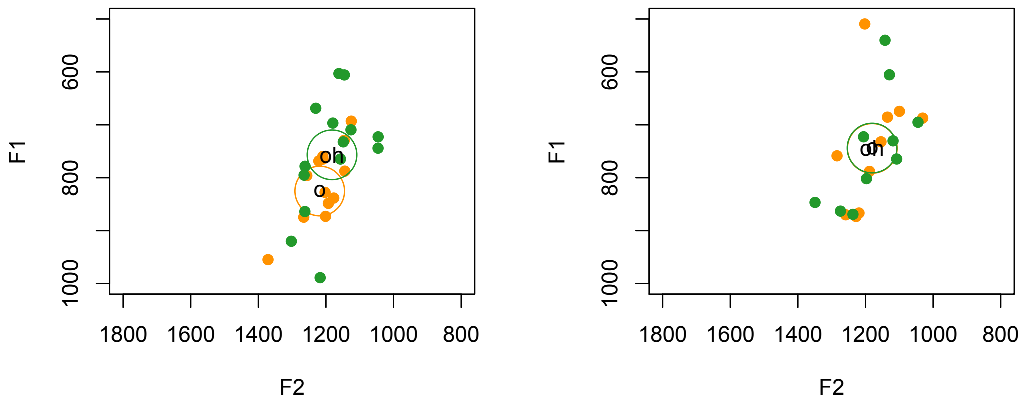
(a) Word List (b) Minimal Pairs

Figure 4.12: JF's (oh) and (o) in read styles



(a) Word List (b) Minimal Pairs

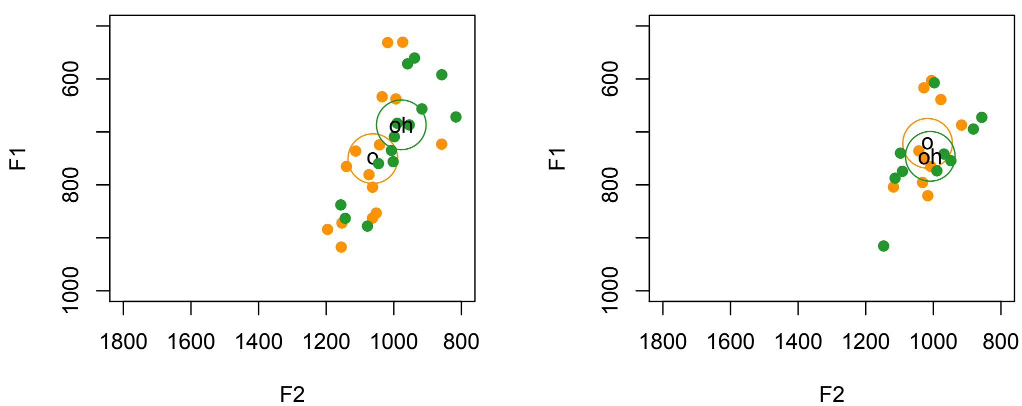
Figure 4.13: CW's (oh) and (o) in read styles



(a) Word List

(b) Minimal Pairs

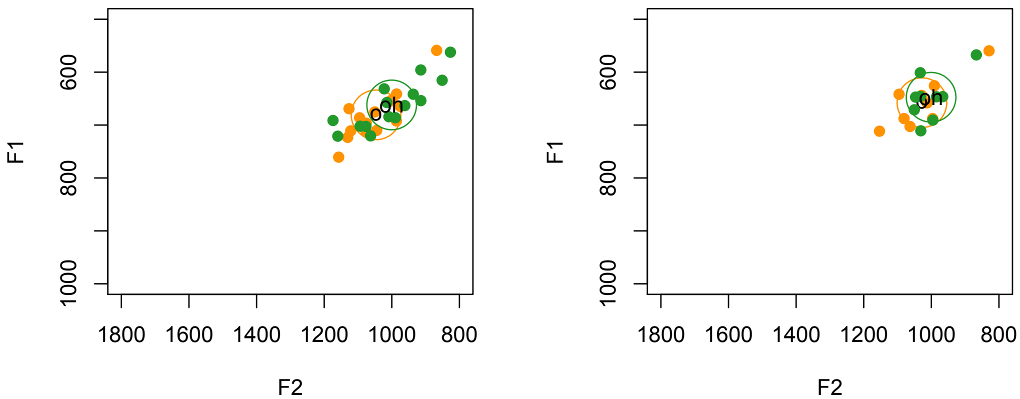
Figure 4.14: NW's (oh) and (o) in read styles



(a) Word List

(b) Minimal Pairs

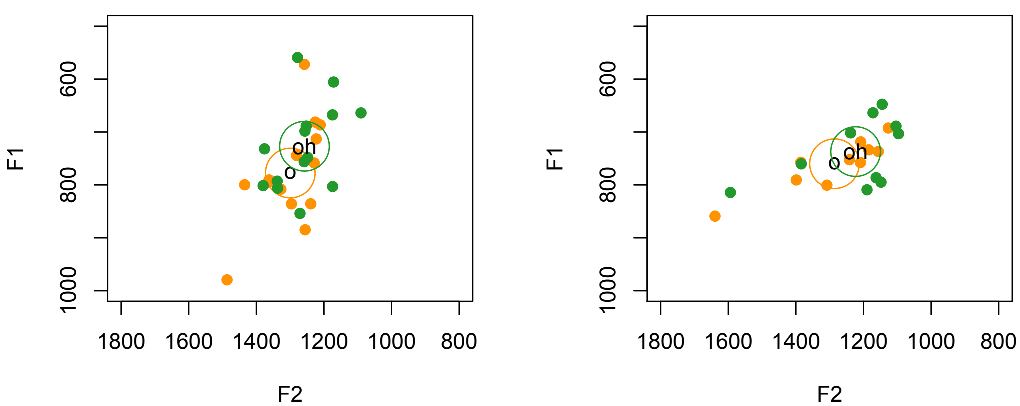
Figure 4.15: LG's (oh) and (o) in read styles



(a) Word List

(b) Minimal Pairs

Figure 4.16: PW's (oh) and (o) in read styles



(a) Word List

(b) Minimal Pairs

Figure 4.17: TM's (oh) and (o) in read styles

	o F1 (Hz)		oh F1 (Hz)					
Speaker	mean	SD	mean	SD	mean difference	t(df)		p
BW	629	28	636	21	-7	t(14) =	1.1928	0.25
CW	848	45	809	63	39	t(13) =	-2.3821	0.03
DB	673	73	659	109	13	t(13) =	-0.5023	0.62
ES	665	85	662	55	2	t(13) =	-0.1493	0.88
EW	589	33	579	35	10	t(14) =	-1.0650	0.3
JF	709	79	690	59	19	t(13) =	-1.3266	0.21
LC	791	54	782	76	9	t(13) =	-0.5638	0.58
LG	750	123	687	137	64	t(14) =	-1.5984	0.13
LW	848	70	825	72	22	t(13) =	-0.8265	0.42
NW	825	91	757	110	68	t(13) =	-2.6927	0.02
PW	681	46	662	46	19	t(14) =	-1.6748	0.12
TM	778	99	727	84	51	t(13) =	-1.8892	0.08

Table 4.7: (o)/(oh) word list production results: F1. Mean and standard deviations for each word class as produced in Word List readings, with paired t test results

Speaker	o F2 (Hz) mean	o F2 (Hz) SD	oh F2 (Hz) mean	oh F2 (Hz) SD	mean difference	t(df)	p
BW	1058	65	1057	47	1	t(14) = -0.0804	0.94
CW	1144	77	1112	130	31	t(13) = -0.9299	0.37
DB	1053	62	1040	78	13	t(13) = -0.5709	0.58
ES	1065	139	1078	54	-13	t(13) = 0.4407	0.67
EW	926	62	933	56	-7	t(14) = 0.5638	0.58
JF	1049	81	1026	74	23	t(13) = -1.6555	0.12
LC	1042	81	1044	96	-2	t(13) = 0.1341	0.9
LG	1062	85	978	104	84	t(14) = -2.650	0.02
LW	1215	99	1206	58	8	t(13) = -0.3287	0.75
NW	1218	70	1182	79	36	t(13) = -1.6947	0.11
PW	1047	77	1000	103	46	t(14) = -2.1575	0.049
TM	1300	85	1258	84	42	t(13) = -2.4267	0.03

Table 4.8: (o)/(oh) word list production results: F2. Mean and standard deviations for each word class as produced in Word List readings, with paired t test results

	o duration (ms)		oh duration (ms)					
Speaker	mean	SD	mean	SD	mean difference	t(df)		p
BW	262	66	250	60	12	t(14) =	-1.0266	0.32
CW	236	82	244	89	-9	t(13) =	0.9361	0.37
DB	225	62	231	77	-6	t(13) =	0.4317	0.67
ES	169	44	198	69	-29	t(13) =	2.3137	0.04
EW	183	34	191	47	-8	t(14) =	1.3686	0.19
JF	164	73	173	63	-9	t(13) =	0.8233	0.43
LC	210	72	206	71	3	t(13) =	-0.2723	0.79
LG	176	55	195	80	-19	t(14) =	1.3296	0.2
LW	166	76	161	59	5	t(13) =	-0.4291	0.67
NW	145	39	184	50	-39	t(13) =	4.4024	<0.001
PW	188	66	182	80	6	t(14) =	-0.4211	0.68
TM	162	56	169	54	-7	t(13) =	0.5812	0.57

Table 4.9: (o)/(oh) word list production results: Duration. Mean and standard deviations for each word class as produced in Word List readings, with paired t test results

4.4 Conversational speech

4.4.1 Acoustic analysis

Every useable token of words from the THOUGHT, CLOTH, LOT and PALM lexical sets were extracted from the recorded interviews and subject to acoustic analysis. "Useable" in this case means any token that showed reasonable formant tracking in Praat; I thus excluded tokens which were produced with excessively creaky or falsetto voice quality, or against background noise. I also excluded tokens which were produced very quickly and were auditorily reduced (i.e. sounded schwa-like); in practice this meant any vowel with a duration of less than 50 milliseconds. I auditorily checked that these performance-related exclusions did not systematically skew the data set.[10] All tokens had primary or secondary stress on the low back vowel.

Tokens were classified as either (o) or (oh), based on how each word is produced in the New York/New Jersey varieties of English which make this distinction. This classification was largely based on my own native speaker intuitions, but was also double-checked with others who share the distinction as well as against sources such as Wells (1982) and Labov et al. (2006). Across all 17 speakers, 2736 conversational tokens of (o) words and 1487 tokens of (oh) were collected for measurement.

Measurements of F1, F2, and F3 were taken for each low back vowel token at the F1 maximum. Measurement points were first marked automatically with a script in Praat, then manually checked for egregious errors and, if necessary, corrected. The F1 and F2 measurements were each then subject to separate statistical analyses.

[10] It was not the case, for example, that the excessively creaky THOUGHT tokens were also the most raised-sounding, or all the falsetto LOTs were the most backed

Formant measurements were not subject to any normalization procedure. Because the main statistical analysis to establish the low back vowel system was carried out for each individual speaker, normalization was not strictly necessary. Moreover, methods of normalization available for analyzing only a portion of the vowel space (such as Bark normalization) are heavily dependent on F3, which is not ideal for vowels which have questionable F3s due to rounding (as low back vowels often do) or having been recorded in sub-laboratory acoustic conditions (as all the interviews in this study were).

4.4.2 Coding

Each token was coded for word class (*o* or *oh*) and phonological context. Four phonological factors were included: *Preceding Place* and *Following Place* each comprise 3 levels: CORONAL, VELAR, and NO LINGUAL; the last level includes all labial and glottal consonants. *Preceding Voice/Manner* includes 9 levels: GLIDE, LATERAL, NASAL, PAUSE, RHOTIC, VOICED FRICATIVE, VOICED STOP, VOICELESS FRICATIVE, and VOICELESS STOP. *Following Voice/Manner* includes 7 levels: LATERAL, NASAL, PAUSE, VOICED FRICATIVE, VOICED STOP, VOICELESS FRICATIVE, and VOICELESS STOP.

4.4.3 Statistical analysis

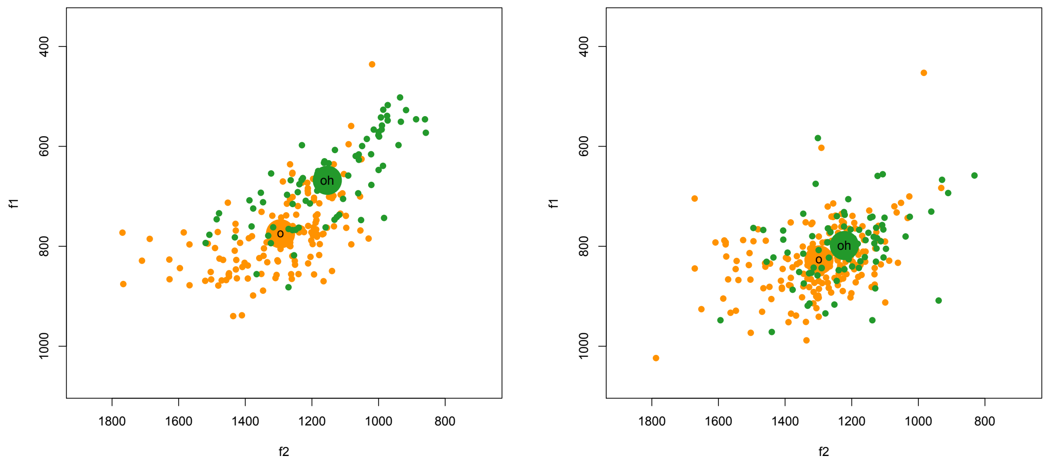
Conversational data is much messier than data elicited using word lists. The linguist has little to no control over the phonological contexts in which the segments of interest occur, and would be lucky to find even a few minimal or near minimal pairs that would allow for a direct comparison of vowel quality in the same context.

Under these circumstances, what is the best method for establishing whether a speaker has a low back vowel distinction?

The simplest approach would be to use t tests to compare the F1 and F2 means of all (o) tokens and all (oh) tokens, in order to determine *whether* there is a significant difference between these groups on either dimension; the actual difference between these means would then indicate *how* distinct these two categories are. For example, consider the conversational low back vowels of Speakers SS and CW, both 54 year-old women who grew up in Montreal and moved to New York at age 27 (Figure 4.18). SS has a significant difference between (o) and (oh) in both F1 mean ($t(139.51)=9.0236$, $p=1.298e\text{-}15$) and F2 mean ($t(149.123)=6.1989$, $p=5.296e\text{-}09$), as does CW (F1: $t(140.663)= 2.9872$, $p= 0.003323$; F2: $t(153.306)=4.0532$, $p= 8.01e\text{-}05$). However, as is apparent from the scatterplots in 4.18, SS has a much greater separation between the means of these vowels than CW, with an Euclidean distance[11] of 177 to CW's 81.

This simple approach is inadequate for at least two reasons. First, phonological context has a strong effect on formant values (e.g. Fant 1960), and so must be taken into account in the analysis. This is especially important in trying to determine whether speakers exhibit a low back vowel contrast, because the relevant word classes are not evenly distributed across phonological contexts (Labov et al. 2006, p. 58-59) For instance, (oh) is much more likely than (o) to occur before tautosyllabic /l/. Following laterals tend to lower the F2 and raise the F1 of preceding vowels, i.e., to apparently back and lower the vowel. If this effect is not corrected for, then the results of the analysis will overestimate the difference between these two word

[11]The Euclidean distance between vowels is calculated by finding the F1 and F2 differences between the vowels, squaring these values, and taking the square root of their sum (i.e. by using the Pythagorean Theorem).



(a) SS (b) CW

Figure 4.18: (o) and (oh) in the conversational speech of SS and CW

classes along the F2 dimension, because (oh) will be disproportionately pulled back in the vowel space by pre-lateral tokens; for the same reason, any difference in the F1 dimension will be underestimated. The analysis needs to establish whether word class membership predicts phonetic realization once phonological context effects have been taken into account.

Second, a simple comparison of means may give an inaccurate impression of contrast (or lack thereof) due to the potentially disproportionate influence of particular lexical items. Certain highly frequent words, such as *not* and *all*, are used many times over the course of an interview by every speaker. Moreover, speakers have their own favorite lexical items. For instance, SS, a psychoanalyst living on the Upper West Side, used *father* 14 times in her interview speech, more than any other speaker uses this word; similarly, TM, who works for the ASPCA,[12] used *dog* or *dogs* an astounding 64 times over the course of our conversation. If there are no

[12]The American Society for the Prevention of Cruelty to Animals

word-specific effects on phonetic realization, then such liberal use of particular lexical items does not pose a problem for analysis; 64 tokens of *dog(s)* can be treated just like any 64-member collection of (oh) tokens. However, if there are universally or idiosyncratically frequent words which behave atypically for their word class, then failing to account for the particular effects of these words may also skew the analysis.[13]

4.4.3.1 Mixed effects models

To overcome both of these issues, I used mixed effects regression analysis implemented using the lmer() function in R (Bates and Sarkar 2008, Pinheiro and Bates 2000, Baayen 2008). Mixed models are a relatively new addition to the sociolinguistic toolbox (see Johnson 2009 for an excellent detailed introduction); they have recently been used in Jaeger and Staum (2005)'s study of *that*-omission and Johnson (2007)'s analysis of survey data pertaining to low back vowel use among schoolchildren in Massachusetts and Rhode Island.

Mixed models incorporate both *fixed* and *random* effects on the realization of dependent variables. *Fixed effects* involve the type of predictors that researchers are typically interested in investigating: in linguistic studies, these might include social factors such as gender or social class. The effects of such factors are expected to be fixed across studies of the same population. If a study is replicated with a different set of speakers, the effect of gender or social class ought to be more or less the same. Individual speakers, however, may to a certain extent vary unpredictably. This is potentially problematic, because the tokens associated with

[13]Of course, words which occur infrequently in a corpus may also behave in phonetically atypical ways. However, the skewing effect of infrequent words will be less serious, as there are, by definition, fewer tokens of them.

any given speaker are also associated with a constant gender and social class; extreme idiosyncratic behavior of specific individuals could therefore be mistaken for more general effects. This problem is dealt with by incorporating Speaker as a *random effect* in the statistical model; random effects are usually not of direct interest, but including them in the model enables a more accurate estimation of the fixed effects.

The nesting relationship which exists between a factor like Speaker and more general social factors like gender and social class may also exist in the linguistic domain. For example, the tokens of any given low back vowel word (such as *dog*) will also have a constant word class - (oh) - and a constant phonological context (post-[d], pre-[g]). To ensure that word-particular patterning does not skew the estimations of the fixed effects of interest (here, phonological context and word class), Word will be included as a random effect in the analyses to follow.

The non-independence of Phonological Context and Word Class seems like a trickier problem. If these factors cannot be perfectly separated, how can the relative contribution of each factor be established? Fortunately, this is not exactly the question I wish to answer. It is a given that phonological context will have some effect on vowel quality; the question is, once these contextual effects have been taken into account, is Word Class *also* a significant predictor of formant values?

Put this way, the problem is easily dealt with given the way that linear models operate in R. Briefly, the ordering of factors in a model matters: if Factor A precedes Factor B in the definition of the model, then Factor A gets first crack at accounting for the variation in the dependent variable.[14] This method suits present

[14]Less briefly: This is due to how Sum of Squares is done in R. R uses a default Type 1: Sequential SS, which means that "the sums of squares shown are the decrease in the residual sums

purposes exactly, as the aim is to factor out phonological effects before estimating the extent to which Word Class explains formant variation.

An advantage of this type of analysis is that it is a fairly conservative method of determining whether contrast exists. There is a danger of making a Type II error: a contrast that really exists in the speech of the speaker may not be detected by this analysis if phonological context and Word Class covary too much, because most of the variation in formant values will be attributed to context. However, if a significant effect of Word Class **is** found, we can be very confident that it is a real effect.

Because mixed models that are fit using lmer() do not provide p-values corresponding to each of the effects,[15] the significance of Word Class was evaluated by taking a model comparison approach, fitting a series of increasingly complex models to the data and checking to see whether adding particular fixed effects ('stepping up') results in a better model. The same procedure was followed for each speaker analysis, for each of F1 and F2:

1. Begin by fitting a null model containing only a random effect of Word. This null model, *m0*, does not include any potentially explanatory fixed effects, returning only effects for each word, which will be equal to the mean formant value for all tokens of that word in the data being analyzed.

2. Fit another model, *m1*, which includes phonological contexts as fixed effects in addition to a random effect of Word. Four phonological fixed effects are included: Preceding Place, Preceding Voice/Manner, Following Place,

of squares resulting from an inclusion of *that term* in the model at *that place* in the sequence." (Venables et al. 2010)

[15]Bates explains why this is the case at the following site: https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html.

and Following Voice/Manner. Using the summary() function, the coefficients associated with each level of these factors can be inspected to see their effects on the relevant formant value.

3. Compare *m0* and *m1* using the anova() function. The result of this comparison answers the following question: Is a model which includes phonological factors better than a model which does not? (The answer, unsurprisingly, is always yes).

4. Fit a third model, *m2*, which includes the phonological fixed effects of *m1*, plus a fixed effect of Word Class, as well as a random effect of Word. Crucially, Word Class is ordered after all of the Phonological factors in the model. Again, the effect sizes associated with all of these factors can be retrieved using summary().

5. Compare *m1* and *m2* using the anova() function. Is a model which includes Word Class in addition to phonological factors better at accounting for the variation in the dependent variable than a model which does not include Word Class?

Two important pieces of information result from this procedure. First, the comparison of *m1* and *m2* reveals whether the model including Word Class is significantly better than a model without this factor: in other words, whether there is evidence that the speaker exhibits a low back vowel difference which is not phonologically predictable, and thus contrastive. Second, the summary of *m2* gives the effect size associated with Word Class; this can be interpreted as an estimate of the real distance in Hz between (o) and (oh) on the relevant dimension, after phonology has been taken into account. These are the results that will be reported for each speaker in the next section.

4.4.4 Conversational speech results

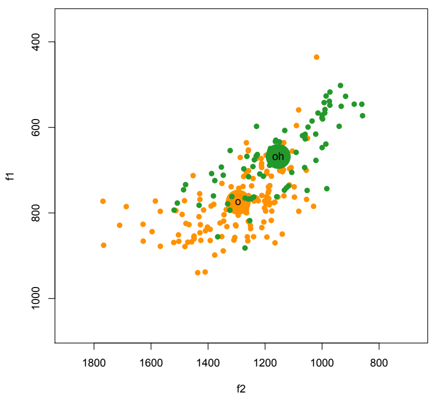
Here I present the results of 17 separate speaker analyses. In typical variationist studies, linguists will combine large numbers of speakers in order to uncover the constraints underlying the community grammar. This practice presupposes that together the speakers a) form a speech community and therefore b) have comparable grammars - that is, are subject to similar linguistic constraints on realization of the variable being studied. These assumptions do not hold for the speakers in this study. While they all share certain social characteristics (all native Canadian, all now living in a similar region), there is no reason to suppose that they will share a system; indeed, the aim of the analysis is to determine whether each speaker continues to exhibit a CE 1P system or has acquired something like a NYAE 2P system, and variation on this point is expected. More generally, the state of having a contrast or not is a characteristic of an individual grammar; it makes little sense to report on whether the group, in aggregate, have acquired a contrast.

That said, the 17 individual speakers fall into a few natural groupings with respect to their results. Word Class is a significant predictor of both F1 and F2 for five speakers: SS, LC, LW, LG, and GH (Fig. 4.19). While the tokens of each word class overlap for these speakers, there is a visible divergence of (o) and (oh) in both height and backness.

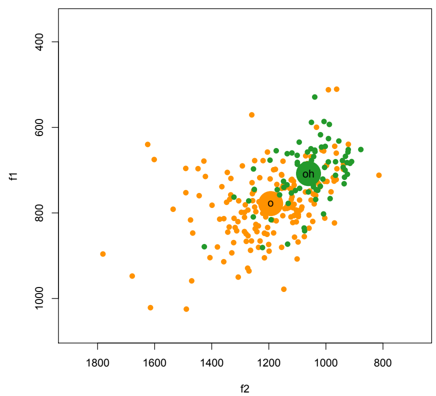
For BK, BW, and JC, there is evidence for contrast only along the F1 dimension (Fig. 4.20). This height difference is clearly visible in the scatterplots of BK and JC's conversational tokens. The plot of BW's vowels does not reveal any obvious height difference; if anything, there appears to be a slight separation along the backness dimension. However, it is important to note that the scatterplots of conversational tokens presented in this section are not adjusted in any way for phonological context. Therefore, it is to be expected that some plots will not seem to match up with the results presented in Table 4.10, which reflect the difference between (o) and (oh) after context has been taken into account.

DB, JF, and EW have contrast only along the F2 dimension, as reflected in the backness differences plotted in (Fig. 4.21).

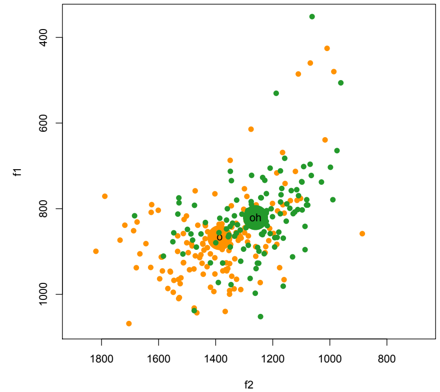
Finally, NW, PW, TM, CW, ES, and VJ have no significant contrast between (o) and (oh) on either dimension (Fig. 4.22). However, even these speakers do not show a complete overlap of the two word classes.



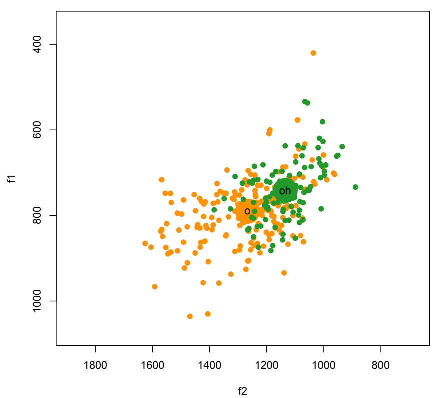
(a) SS



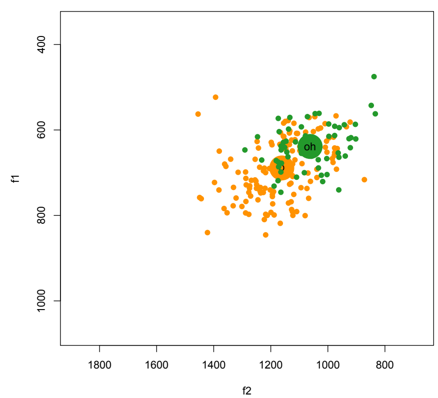
(b) LC



(c) LW

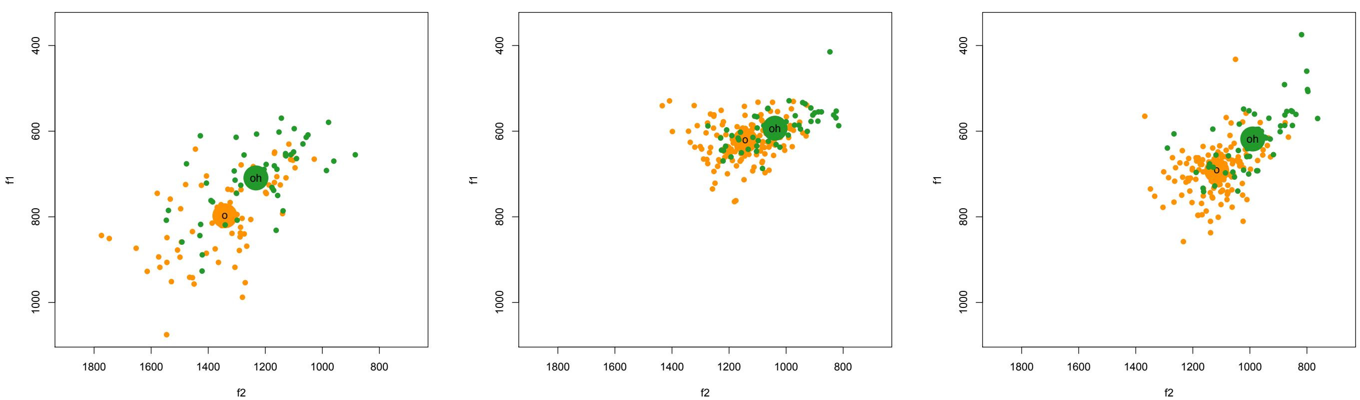


(d) LG



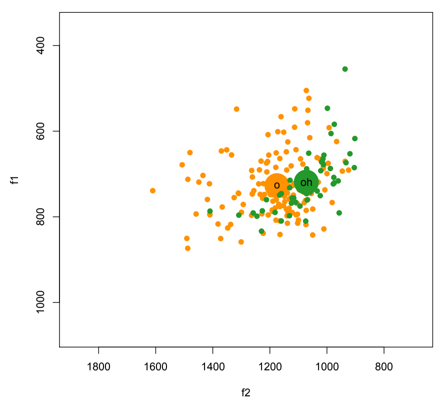
(e) GH

Figure 4.19: (o) and (oh) in conversational speech: Speakers with contrast in F1 and F2

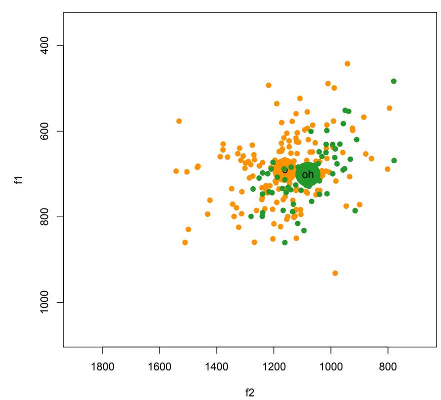


(a) BK (b) BW (c) JC

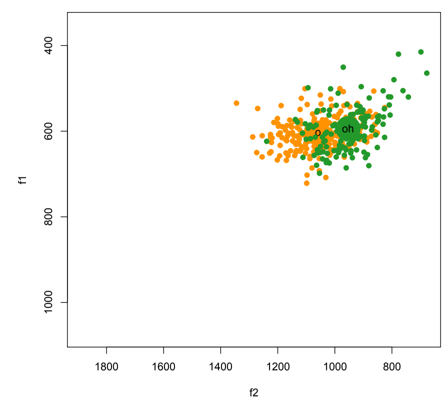
Figure 4.20: (o) and (oh) in conversational speech: Speakers with contrast in F1 only



(a) DB

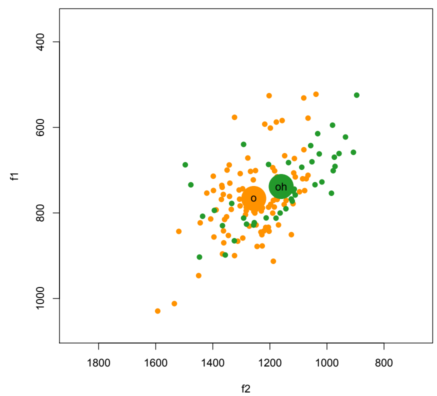


(b) JF

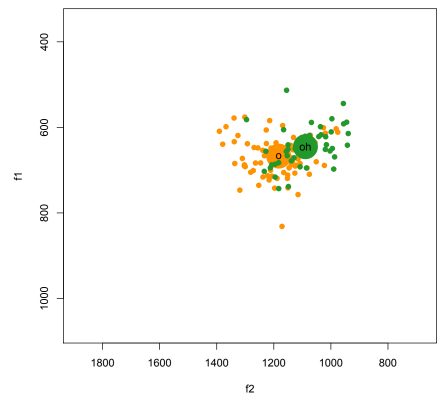


(c) EW

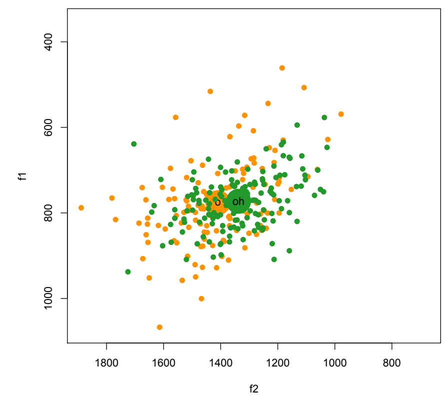
Figure 4.21: (o) and (oh) in conversational speech: Speakers with contrast in F2 only



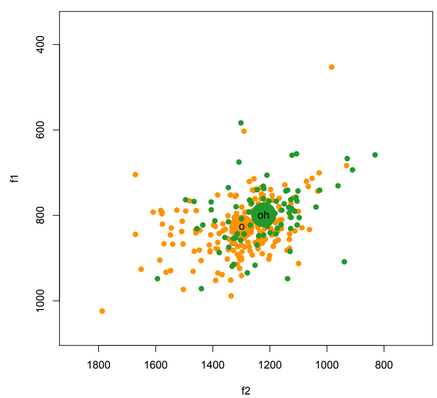
(a) NW



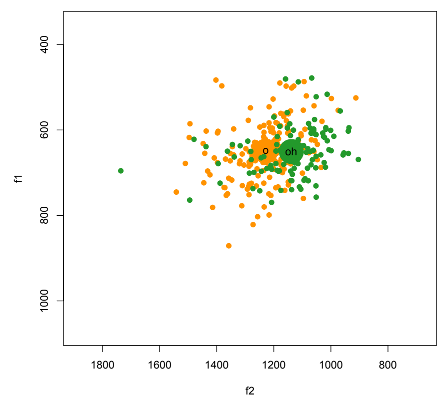
(b) PW



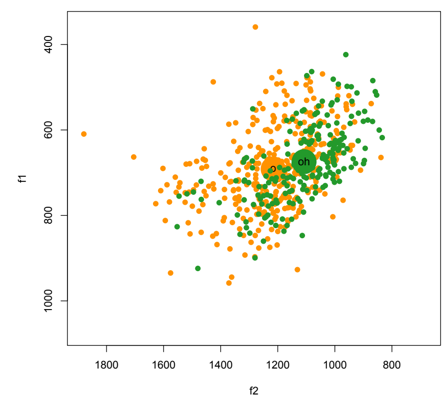
(c) TM



(d) CW



(e) ES



(f) VJ

Figure 4.22: (o) and (oh) in conversational speech: Speakers with no contrast

	F1		F2		Euclidean
Speaker	Word Class effect?	Effect Size (Hz)	Word Class effect?	Effect Size (Hz)	distance
SS	yes	81	yes	83	116
LC	yes	38	yes	73	82
LG	yes	36	yes	63	73
LW	yes	44	yes	57	72
GH	yes	44	yes	36	57
BK	yes	66	no	19	69
JC	yes	42	no	32	53
BW	yes	29	no	25	38
DB	no	22	yes	62	66
JF	no	13	yes	56	57
EW	no	10	yes	53	54
NW	no	15	no	37	40
TM	no	15	no	28	32
PW	no	13	no	28	31
CW	no	8	no	9	12
VJ	no	8	no	9	12
ES	no	3	no	0	3

Table 4.10: (o)/(oh) conversational results. The results of the model comparison for each formant are given for each speaker: first, whether the model which includes Word Class as a predictor in additional to phonological factors is significantly better than a model which does not include this predictor, and second, the effect size associated with Word Class. The rightmost column contains the Euclidean distance derived from the F1 and F2 effect sizes.

Based on the analysis described above, eleven speakers show evidence of a distinction between (o) and (oh) in their conversational speech. These results are summarized numerically in Table 4.10, which presents both the effect sizes associated with each phonetic dimension and whether this difference was found to be significant. Note that larger effect sizes tend to be associated with significance, but this is not always the case. Along the F1 dimension, effect sizes of 29Hz and higher are significant, while lower effects are not. Along the F2 dimension, however, the cutoff seems to be around 36 or 37Hz (the former is significant for GH, while the latter does not reach significance for NW; this difference is likely attributable to the greater degree of overlap found in NW's realization of the two word classes). Again, the effect size for each dimension can be interpreted as the real difference between the two categories along this dimension; that is, the difference that would exist for words from these classes if phonological context were controlled for.

From these numbers, it is possible to calculate the real Euclidean distance between (o)and (oh) in the two-dimensional vowel space; these figures can be found in the last column of Table 4.10, and are visualized in Figure 4.23, which plots the relative effect sizes across speakers. Figure 4.23 depicts a sort of idealized vowel space, with the placement of each speaker's (o) and (oh) in this space reflecting the magnitude of the difference between them. Thus speakers with a large difference along both dimensions (SS, LW, LG, LC, GH) are plotted farther away from the origin, while speakers with very small effect sizes (ES, CW, VJ) are clustered near the origin.

Two additional points arise from these results. First, while 11 of 17 speakers show a significant difference along at least one dimension, there is wide variation in terms of how this difference is realized; SS, the speaker with the most robust

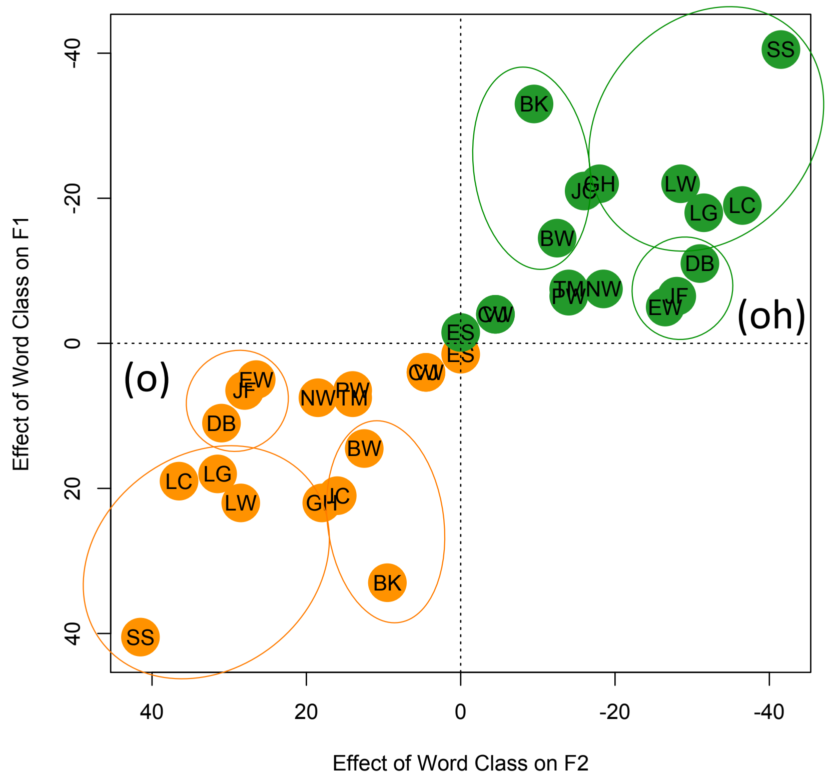


Figure 4.23: Real distance between (o) and (oh) by speaker, based on the effects of Word Class on F1 and F2

distinction, has a Euclidean distance of 116Hz between (o) and (oh), while BW has a distance of 38Hz. Second, even among speakers with no significant difference along either dimension, both effects trend in the same way: (oh) is consistently higher (and usually backer) than (o).

4.4.4.1 Frequency effects

The analysis thus far has established that natively 1P speakers are capable of acquiring a distinction between (o) and (oh). In addition, this distinction seems to be acquired in a phonetically gradual manner: there is no clear separation of two categories, but much overlap in the vowel space. This section presents the results of an analysis to determine whether this distinction is also acquired in a lexically gradual manner, by testing for frequency effects on the realization of (o) and (oh).

Two issues arise here. First, it is necessary to determine the right measure of frequency. Various corpora exist from which frequency counts can be obtained, but these fall short in various ways: many are based on written speech (e.g. CELEX (Baayen et al. 1993)), some are based on dialects of English which are not spoken by the speakers in this study (e.g. the British National Corpus), and others are simply out of date (e.g. the Brown Corpus (Kučera and Francis 1967)). Moreover, as noted in Section 4.4.3, while certain words occur with high frequency in all 17 interviews, reflecting the commonality of these words in the linguistic input of all speakers, other words are idiosyncratically frequent, in ways which seem to reflect the individual lived experience - and likely linguistic input - of each speaker. For this reason, I used a speaker-internal measure of frequency, counting up the number of times each word appears within a given speaker's interview. For example, the word *problem* may be coded as frequency 6 for a speaker who uses that word 6 times, but as 2 for a speaker who uses it only twice over the course of an interview. Frequency counts here are simply raw counts over the course of the interview. However, as all interviews were of roughly comparably duration (1.5 hrs), the counts should likewise be roughly comparable across speakers.

The second issue is that there are not enough tokens from each speaker to examine frequency effects at the speaker level, especially once phonological effects and word class have been taken into account. In addition, any given word within a single speaker's data will always have both a particular phonological context and a particular interview frequency. To address these issue, I pooled all speakers for the frequency analysis. This approach both increased the amount of data in the analysis and to some extent decoupled phonological context and frequency: given the speaker-specific frequency coding described above, words may vary in frequency coding across speakers. In order to pool the unnormalized formant data, it was necessary to correct for the gross formant magnitude differences across speakers; this was accomplished by including a Speaker random effect in the models described in the analysis below.

The aim of the following statistical analysis is to determine whether there are frequency effects on the realization of (o) and (oh). Usage-based models predict that high frequency lexical items should be more advanced with respect to phonetic changes; in the present case, this means that high frequency low back vowel words should be the first to show signs of shift towards NYaE. Importantly, this change involves word class divergence: (o), when it moves, shifts fronter and lower in the vowel space, while (oh) shifts higher and backer in this space. The effect of frequency is therefore expected to interact with word class. Higher frequency (o) words are predicted to have higher F1 and F2 than low frequency (o) items, while high frequency (oh) words ought to have lower F1 and F2 than low frequency items in that category.

To test each of these predictions, separate analyses of F1 and F2 were completed for each of (o) and (oh). Again, a step-up model comparison procedure was used. For each analysis, the crucial comparison was between a model, m1, which includes only fixed phonological effects and a random effect of Speaker, and m2, a model which includes the same fixed phonological effects, the random effect of speaker,

	Effect of Frequency (Hz/Count)	p
(oh) F1	-0.38	0.019
(oh) F2	(-0.03)	1
(o) F1	0.52	0.008
(o) F2	1.72	<0.001

Table 4.11: Frequency effects on F1 and F2 for (o)/(oh) in conversation

and a fixed effect of frequency. Results of the 4 pooled analyses are summarized in Table 4.11.

First consider (oh). Frequency has a negative effect on both formants, though this is only significant on the F1 dimension: higher frequency is associated with lower F1 values even after phonological factors have been taken into account. This means that high frequency (oh) words are realized in a higher position in the vowel space than low frequency words of this word class. For (o), however, frequency has a significant positive effect on both formants: higher frequency (o)s are associated with higher F1 and higher F2 values; that is, high frequency (o) words are realized lower and fronter than their low frequency counterparts. In short, the predictions outlined earlier in the section are borne out by the conversational data.

4.5 Discussion of results

This chapter examined the realization of low back vowels in the speech of 17 native Canadians who have moved to the New York region. While CE is characterized by a merger between (o) and (oh), NYAE distinguishes these two word classes, with (oh) being realized in a higher and backer region of the vowel space than (o). Two questions were asked and answered in these pages. First, is there evidence that any of the speakers in the sample have acquired a contrast between (o) and (oh)? Second, how is this contrast acquired: to what extent is it lexically gradual, and what is the magnitude of the resulting phonetic change?

The answer to the first question is a yes, though perhaps a qualified one. It must be noted that to a certain extent, these speakers show remarkable stability in their low back vowel system. This is most clearly evident in the Minimal Pair results: nearly all speakers are merged in production and perception in this context. Where speakers do make a significant distinction between (oh) and (o), the phonetic difference is quite subtle compared with the robust distinction made in NYAE.

That said, 11 of the 17 speakers do show evidence of having acquired a distinction between (o) and (oh), on at least one phonetic dimension. That is, these speakers show phonetic variation in these vowels that cannot be attributed to phonological context alone, but can be at least partially explained by word class membership in the ambient dialect. This change, where is has occurred, seems to be phonetically gradual: there remains extensive overlap between the two word classes, indicating a gradient separation.

A possible generative account of these results might be that these speakers have managed to change their underlying forms for some relevant lexical items to reflect the contrast in their new dialect: words such as *cot* and *caught* which had previously been represented identically as [kɑt] and [kɑt] are now stored as [kɑt] and [kɔt], respectively. However, such an account brings with it the puzzling implication that these newly contrastive representations must then be submitted to phonetic implementation rules which all but neutralize the distinction in surface forms! In a usage-based account, however, the subtlety of the surface distinction is easily accommodated, and indeed predicted: contrast is not achieved in a featural quantum leap, but gradually, via the addition of exemplars at the word level, which gradually lead to a more general divergence at the word class level.

Further support for a usage-based account comes from the frequency effects observed in this data. High frequency (oh) words are higher than other (oh) words, while high frequency (o) words are lower and fronter, indicating that high frequency items are on the vanguard of divergent shift within their respective word classes in the low back vowel spaces of these speakers. These facts indicate a lexically gradual shift towards the new variety: speakers hear high frequency words more often, meaning that they acquire new dialect exemplars of these words at a greater rate, which results in the representations (and later productions) of these words shifting before those of less frequent words. These results are difficult to accommodate within the generative account; the best it can do is posit lexical exceptions which generate these results, but in such an account the fact that these exceptions are structured in terms of frequency would be mere coincidence.

Finally, it is worth nothing that the frequency effects reported for conversational speech are consistent with the overall patterns of style shift shown by speakers across Word List and Minimal Pair tokens. While higher frequency items of both word classes are more advanced in the shift towards NYaE in conversation, there is an asymmetry in the magnitude of these effects: high frequency (o) items are more advanced with respect to frontness and height, while high frequency (oh) items differ only in height; moreover, the effects are greater for (o), indicating that this word class is undergoing more shift. A similar pattern occurs in the read styles: for speakers who separate these vowels in the Word List context, it is (o) which shows the greatest shift from Minimal Pair productions.

4.5.1 Variation across styles

As noted above, the speakers in this study show context-dependent variation in their realization of the low back vowel contrast. While 11 speakers make a significant distinction between (o) and (oh) in their conversational speech, none of these speakers exhibit that distinction in Minimal Pair speech.[16]

This is strange behavior for a Minimal Pair task. Minimal pair lists highlight possible contrasts, and are thus the context in which contrasts - even marginal ones - are most likely to surface. In Labov (1966)'s study of (r) on the Lower East Side, for example, speakers contrasted word pairs like *sauce/source* most consistently in the Minimal Pair context, using more coda (r) in this style versus the connected speech styles. Even in cases of near-merger, where speakers do not themselves perceive the difference in their speech, the marginal contrast will reveal itself in minimal pair tests (Labov 1991). The Canadians in this study, however, behave in the opposite way: the marginal distinction in their conversational speech is essentially eradicated in just the context in which it should be most likely to appear.

An explanation for this patterning may come from considering just what Minimal Pair tasks are meant to elicit. (Labov 1966, p. 152) sets Minimal Pair tasks (along with Word Lists) apart from the connected speech styles he analyzes, noting that the citation styles are better taken as an indication of "phonic intention, illustrating the norms of the speaker, in part, rather than a reliable indication of performance." In the case of the New Yorkers Labov interviewed, the norm which was illustrated in Minimal Pair speech was (r)-fulness; this reflected the

[16]JC, the one speaker who shows a significant distinction in Minimal Pair tokens, does not make the same distinction in conversational speech: his conversational vowels differ along the F1 dimension, while his Minimal Pairs differ only along the F2 dimension.

local change in progress towards the wider norm of realizing coda (r). Labov's speakers may not have consistently produced (r) in their connected speech, but at some level they knew that they *should* do so.

The expatriate Canadians in this study find themselves in a very different social context. They are not natives of a speech community undergoing change, but newcomers to a community with stable, though different, norms. However, these new norms do not seem to be adopted as such by the mobile speakers, *even though* they show measurable effects on the speech of these speakers. Instead, it seems that the Canadian speakers maintain their D1 norms for low back vowel realization.

Apart from the linguistic conclusions discussed above, this chapter presented two types of results for each speaker: *whether* that speaker makes a low back vowel distinction, and a measure of the *magnitude* that distinction. In Chapter 6, this information will be used to compare speakers' realization of (o)/(oh) with their use of Canadian Raising. Chapter 7 will examine the extent to which either of these measures correlates with various social factors. Bur first, Chapter 5 will present the results of the Canadian Raising analyses.

CHAPTER 5

CANADIAN RAISING

As discussed in Chapter 3, phonological generalizations have a very different status in generative theory versus usage-based theories. In generative phonology, generalizations about the patterning of sounds are formalized as rules or constraints, which map an input form onto an output form, and are crucially separate from either of these representations. In usage-based theories, generalizations are epiphenomenal: they do not act upon lexical representations or map one representation to another, but arise out of these representations.

This differing theoretical status entails different mechanisms underlying how generalizations may change over the lifespan. The rule itself is a potential locus of change in generative theory: it can be altered, deleted, or made variable. In usage-based theories, lexical representations are the primary locus of change; however, many lexical changes in the same direction may accumulate and "trickle up" to result in a change in the generalization that is derived from the collection of exemplars. Each theory thus implies a particular course of change across the lifespan if speakers are to alter a phonological generalization. If generalizations are really separately alterable rules/constraints, then we expect changes in generalizations to affect all lexical items containing the relevant segments in the relevant contexts. If, however, generalizations arise from lexical representations, then we expect to see changes occuring on a lexically gradual basis.

This chapter will examine how the speakers in this study have changed with respect to a particularly salient phonological generalization: the raising of (aw) in pre-voiceless contexts.

5.0 The research questions

This chapter addresses two main questions. First, do the Canadians in this study show evidence of having maintained a Canadian Raising rule, or have they lost this generalization after exposure to NYaE? This question really has two parts: is there a measurable difference in vowel height between (aw) tokens which occur before voiceless consonants and those found elsewhere, and - importantly - is this difference of a large enough magnitude to say that the speakers are exhibiting Canadian Raising?

Second, to the extent that Raising has been lost or attenuated in the speech of these speakers, how has this change occurred: is the change lexically gradient or abrupt?

To answer these questions, tokens of (aw) produced in conversational context were examined both for differences between raising and nonraising contexts, as well as possible frequency effects on realization of this vowel.[1]

[1] Realization of (aw) in read styles was not analyzed: no Minimal Pair productions or judgments were collected for this variable, as it represents a case of allophony, not contrast; also, only a small number of tokens were collected in Word List style.

5.1 Acoustic analysis

Every useable token of words from the MOUTH lexical set was extracted from the conversational portion of each interview and subject to acoustic analysis. Again, "useable" means any token that proved amenable to acoustic analysis, as described in Chapter 4. Very short (<80ms) and auditorily reduced tokens were excluded. Tokens of (aw) in which the diphthong appeared before a nasal were also excluded.[2] Across all 17 speakers, 1210 conversational tokens of (aw) words were collected for measurement.

An F1 measurement was taken for each diphthong at the nuclear F1 maximum. Measurement points were marked automatically with a Praat script, then manually checked for errors and, if necessary, corrected. Again, formant measurements were not subject to any normalization procedure.

5.2 Statistical analysis

A speaker who exhibits Canadian Raising produces two audibly and measurably different allophones of (aw), one which occurs before voiceless consonants and one which occurs elsewhere. Pre-voiceless tokens of (aw) (which I will abbreviate as (aw)T throughout this chapter) are realized with a higher-sounding nucleus, and have on average lower first formants than tokens of (aw) in other contexts (henceforth (aw)O).

One approach to investigating whether a speaker exhibits Canadian Raising is simply to compare the mean F1 of her (aw)T tokens with that of her (aw)O tokens. For example, there is much separation between (aw)T and (aw)O in the

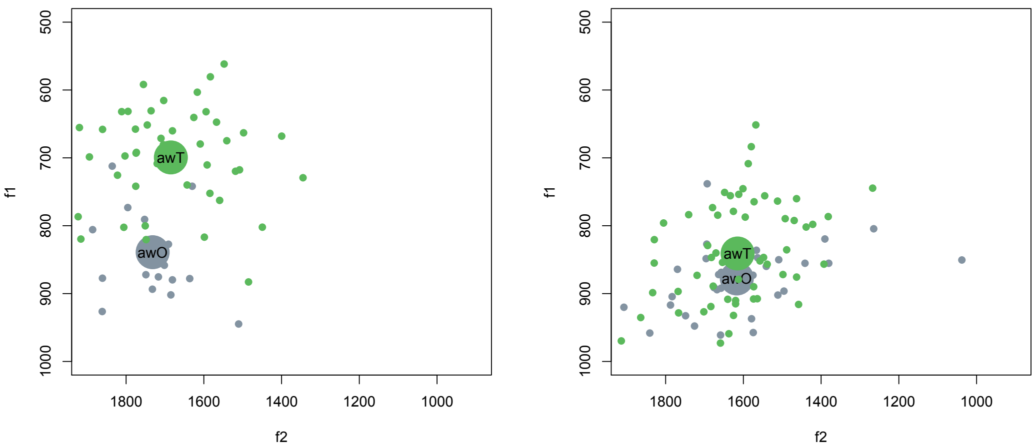
[2]Pre-nasal (aw) tends to be fronted compared to pre-oral (aw) in CE (Boberg 2008).

conversational speech of SS (Fig. 5.1a): the difference in mean F1 is 140 Hz, well above the 60Hz threshold used by Labov et al. (2006) to categorize speakers as exhibiting raising. CW, meanwhile, shows only a 50Hz difference between the groups (Fig. 5.1b), which is below Labov et al. (2006)'s threshold. A reasonable conclusion to draw from this comparison is that SS continues to exhibit robust Canadian Raising, while CW does not.

T tests show that the difference between (aw)T and (aw)O is significant for both SS (t(27.662)=6.8362, p= 2.123e-07) and CW (t(91)=3.2108, p= 0.00183). However, in this case of allophony one must be wary of accepting just any significant difference between these two groups as indicating the presence of Canadian Raising. Any speaker of English is likely to have some degree of raising in a pre-voiceless environment, as this is a shortening environment; this difference may even be significant. However, it will not always correspond to the percept of an auditorily raised (aw)T. For this reason, statistical significance will be considered a necessary but not sufficient condition for establishing that a speaker shows Canadian Raising; the size of the difference will also be noted and compared to the 60Hz threshold set by Labov et al. (2006).

Unfortunately, the simple analysis described above will paint a distorted picture of what is happening with the Canadian Raising. In this case, the lack of phonological control is less of a concern, because phonological context - specifically, following voicing - is the very thing that defines the allophonic classes to be compared. Moreover, the range of possible following environments is much more restricted, because (aw) in English appears only syllable-finally or before a coronal consonant[3].

[3]Except in proper names, e.g. *Houk* and *Laub*.



(a) SS (b) CW

Figure 5.1: (aw)T and (aw)O in the conversational speech of SS and CW

word	n
about	312
out	318
other raising context word (e.g. *doubt, south*)	208
nonraising context word (e.g. *crowd, loud*)	372
Total	1210

Table 5.1: Distribution of (aw) tokens in the whole corpus

A more serious challenge for the analysis of this dataset is the potential for word-specific skewing effects. As noted above, every useable token of (aw) was extracted from the 17 interviews. As it turns out, about half of these are tokens of the words *out* and *about*, which occur extremely frequently in the dataset (Table 5.1)

If these highly frequent words are folded in with other less frequent (aw)T words in the analysis, the conclusions regarding whether a speaker exhibits Raising are going to be overwhelmingly affected by the behavior of these two items. This imbalance needs to be addressed in the statistical model in some way. A mixed model with a random effect of Word, as was used for the low back vowel analysis, is one possible way to do this. However, there are two reasons why I will not take this approach. First, in this particular dataset, *about* and *out* are essentially unique in appearing with high frequency in everyone's speech; the words *how* and *now* are also common, but much less so (and not in all speakers). Most words are represented by 1-3 tokens in any given speaker's interview speech. Therefore, using mixed effects modeling will not accomplish much beyond isolating the effects of *about* and *out*. The same result can be achieved in an ANOVA by simply separating *about* and *out* from the rest of the raising context words and testing all of these against the nonraising context group.

There is an additional reason why it is desirable to separate *out* and *about* from other less frequent raising words. In addition to being highly frequent words, they are highly salient words with respect to Canadian Raising: when speakers (both in this study and outside of it) are asked about the linguistic features which distinguish Canadian English, they invariably mention the word *about* and/or the phrase *out and about*. It is reasonable, therefore, to ask whether these words might pattern differently from other, less salient raising context words. It might be the case, for example, that heightened awareness of the pronunciation of these words might cause speakers to attenuate their use of Canadian Raising in *about* and *out* to a greater extent than in other words which might be subject to raising.

For these reasons, the analyses that follow will compare F1 values in *about*, *out*, other raising context words, and nonraising context words. This approach will reveal whether each speaker shows Canadian Raising in general, as well as in the shibboleth words.

To compare these groups, a series of ANOVA were run in R using *treatment contrasts*. In this contrasts setting, one level of the factor group that is being tested is taken to be the default level, against which other levels of the factor are compared. As the name suggests, this is an appropriate type of analysis for studies which involve a control group and one or more treatment groups, and the main question of interest is whether (and how) the treatment groups differ from the control group. This manner of contrasting is usually not appropriate linguistic analysis, because linguistic factors typically do not have a clear default level. [4]

For the (aw) raising data, however, treatment contrasts are a useful means of determining the degree of raising exhibited in each context by the speakers in this study. There are four groups of tokens to be compared: *about*, *out*, other (aw)T tokens, and (aw)O - and the main question of interest is how the first three groups differ from the last. An ANOVA with treatment contrasts was carried out for each speaker, with Lexical Context as a predictor variable comprising 4 levels (ABOUT, OUT, AWT, and AWO) and AWO set as the default level in the analysis. The result of this analysis for each speaker was an ANOVA table including effect sizes (in Hz) for each of ABOUT, OUT, and AWT, indicating how much each of these groups differs from AWO. A series of pairwise t tests with a Holm correction was then used to assess the differences between all of these groups.

[4] For instance, when testing the effect of following consonant place on formant measurements, there is no clear sense in which one particular place is the default against which the effects of other levels should be compared. This is why, in the analysis of low back vowel raising in Chapter 4, sum contrasts were used. The use of treatment contrasts for such factors is not *wrong*, but it does complicate the interpretation of the analysis: instead of returning effects for each factor that can be evaluated in relation to an abstract group mean, the effect of any given factor must be interpreted in relation to the default level.

Speaker	AWT	OUT	ABOUT
LC	-171	-132	-193
DB	-105	-131	-128
PW	-78	-87	-109
GH	-93	-114	-143
LG	-66	-107	-140
EW	-81	-73	-132
SS	-99	-97	-174

Table 5.2: Pattern 1: General Raisers. F1 effect sizes (in Hz) associated with each Raising group, with AWO context taken as the baseline.

5.3 Conversational speech results

This section presents the results of the individual speaker analyses. Though there is variation across speakers in terms of where the four Lexical Contexts sit in the vowel space and the relationships between these groups, the speakers do exhibit two major patterns.

5.3.1 Pattern 1: General Raisers

Seven speakers exhibit a pattern of general raising: that is, each of the raising context groups ABOUT, OUT, and AWT are found to be significantly higher than AWO. Moreover, in each case, the difference exceeds the 60Hz threshold. Table 5.2 summarizes the effect sizes associated with each Lexical Context for Pattern 1 speakers. AWO is the baseline against which these numbers should be compared; for example, the F1 of LC's AWT group is 171Hz lower than the AWO group, while her ABOUT group is 193Hz lower.

For LC, DB, and PW, ABOUT, OUT, and AWT are each significantly different from AWO according to the post hoc pairwise tests (Fig. 5.2). For these three speakers, there are no significant differences among the three raising contexts.

However, in each case, ABOUT seems to be higher than the AWT group; for LC and PW, ABOUT is also higher than OUT.

GH and LG are also General Raisers, but they show a finer-grained distinction among the raising contexts than the first three speakers (Fig. 5.3). For these two, ABOUT is significantly higher than AWT. The status of OUT is less clear; it is not significantly different from ABOUT or AWT, being located approximately midway between these two groups in each case.

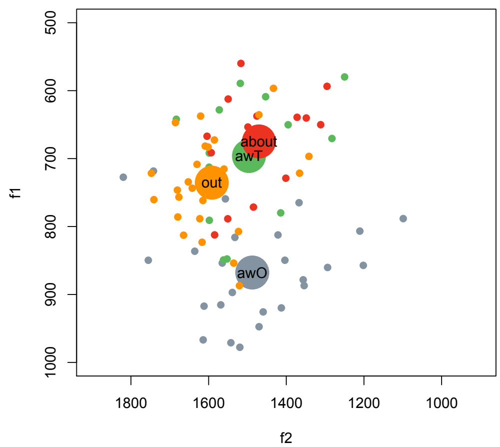
SS and EW show a clearer two-way grouping among the raising contexts: for both of these speakers, ABOUT is significantly higher than OUT, which patterns with the AWT group (Fig. 5.4).

While there is variation among these speakers in terms of which raising contexts are significantly different from each other, an inspection of the effect sizes and the scatterplots in Fig. 5.2, 5.3, and 5.4 reveal the following generalization: across speakers, ABOUT tends to be the the most raised of the raising context groups, AWT is the least raised, and OUT is often somewhere in between.

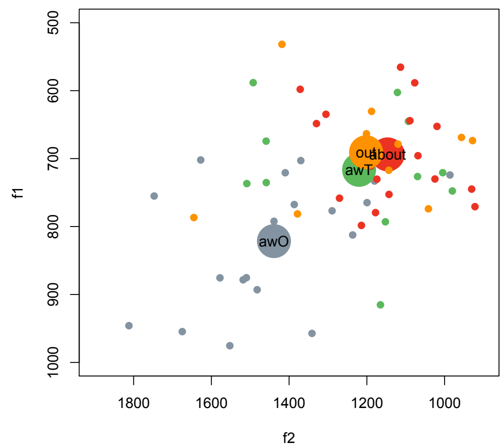
5.3.2 Pattern 2: *out* and *about* Raisers

Eight speakers show no significant difference between AWO and AWT according to post hoc comparisons; moreover, the difference between these two lexical contexts for these speakers is in most cases below the 60Hz threshold. However, these speakers do show significant raising of one or both of ABOUT and OUT. Table 5.3 summarizes the F1 effect sizes associated with each Raising context for Pattern 2 speakers.

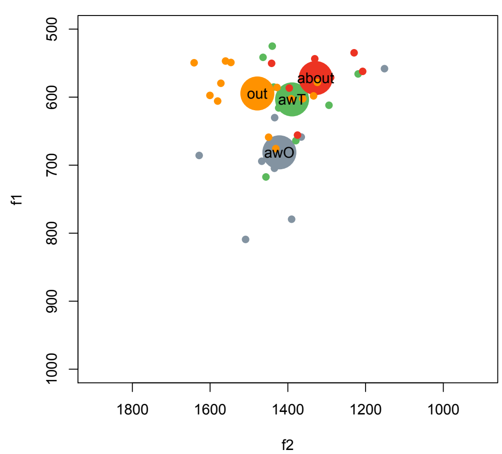
For half of this group - TM, BW, LW, and VJ - ABOUT and OUT are both significantly higher than the AWO group. In the speech of VJ and BW, OUT and



(a) LC

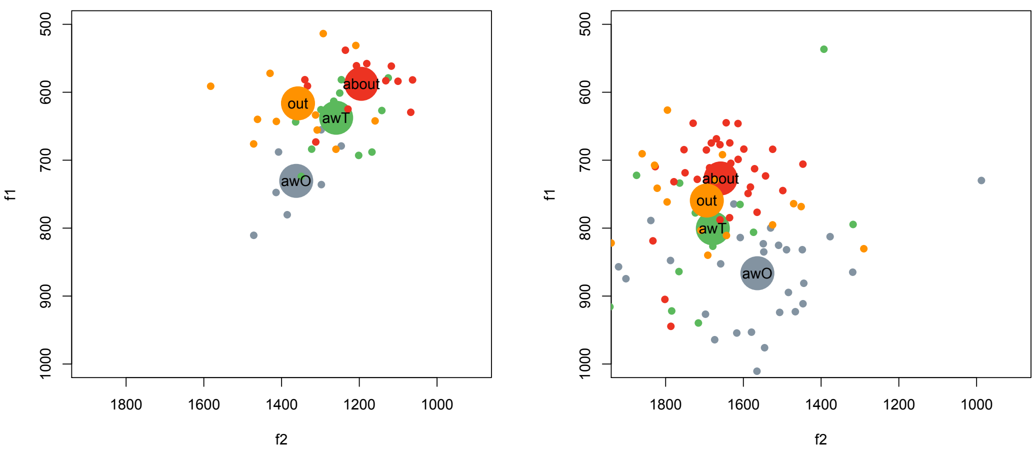


(b) DB



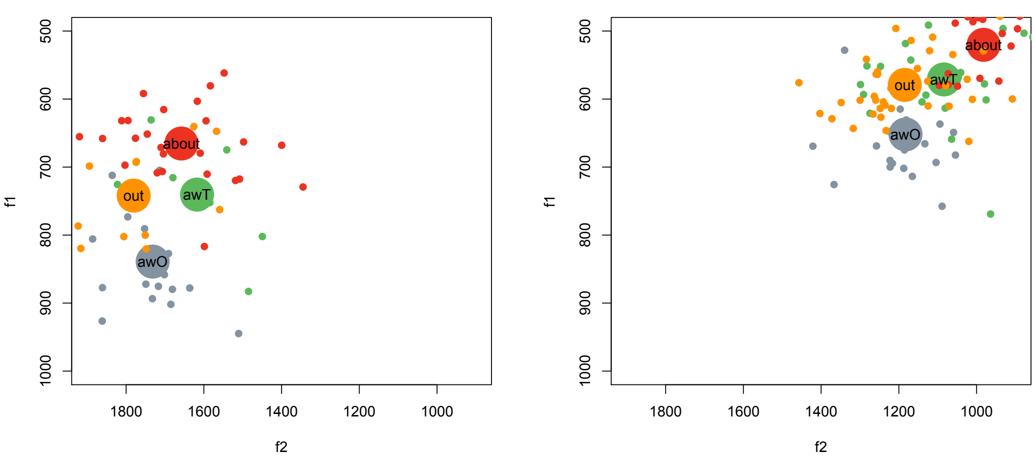
(c) PW

Figure 5.2: Pattern 1: LC, DB and PW raise ABOUT, OUT, and AWT compared to AWO words. There is no significant difference between the three raising contexts.



(a) GH (b) LG

Figure 5.3: Pattern 1: GH and LG raise *about*, *out*, and AWT words compared to AWO words; ABOUT is also significantly higher than AWT.



(a) SS (b) EW

Figure 5.4: Pattern 1: SS and EW raise *about*, *out*, and AWT words compared to AWO words; *about* is significantly higher than *out* and the other raising words.

Speaker	AWT	OUT	ABOUT
BW	(-49)	-80	-99
VJ	(-3)	-63	-67
TM	(-65)	-111	-133
LW	(-19)	-79	-180
NW	(-62)	(-71)	-143
CW	(-24)	(-20)	-55
ES	(-9)	(-38)	-85
JF	(0)	(-61)	-82

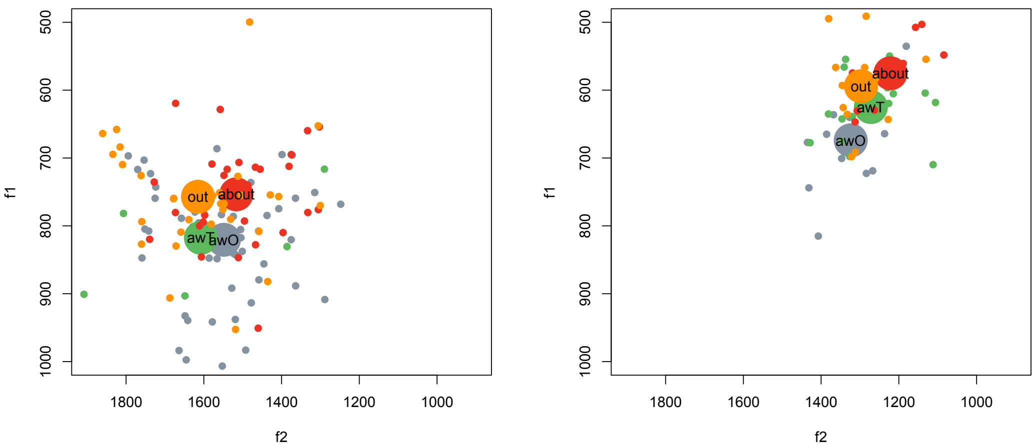
Table 5.3: Pattern 2: OUT and ABOUT Raisers. F1 Effect sizes associated with each Raising context, with AWO context taken as the baseline. Effects in () are not significant, according to the post hoc pairwise tests

ABOUT are not significantly different from each other, nor are they significantly different from the AWT group (Fig. 5.5).

TM and LW also raise both ABOUT and OUT. For these speakers, ABOUT is also significantly higher than AWT. LW's ABOUT is also different from her OUT. Interestingly, both of these speakers also have a visibly backed ABOUT, further distinguishing this item from the other raising contexts. (5.6)

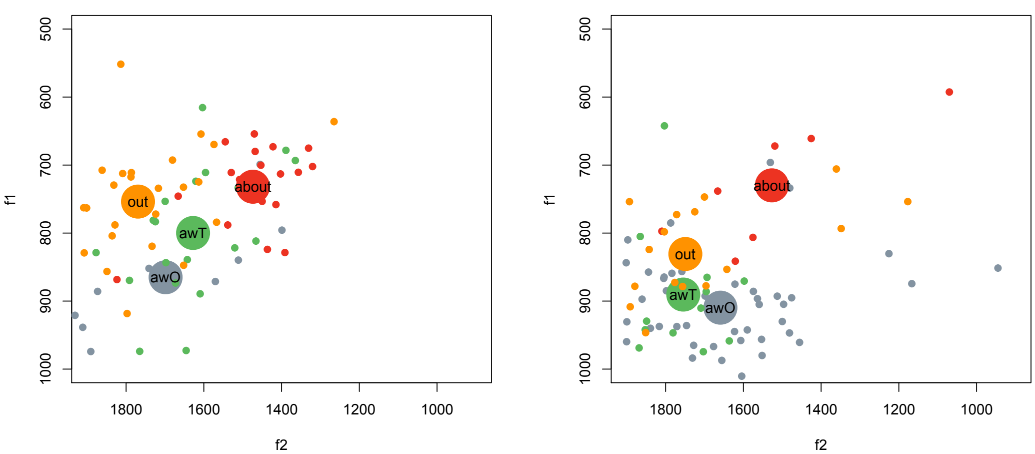
The other 4 speakers in this group show significant raising only in ABOUT. NW and CW's ABOUT is significantly different from AWO, though not significantly different from the AWT contexts, which fall between ABOUT and the AWO group. (Fig. 5.7) In these speakers as well, ABOUT is backed compared to the other groups.

ES and JF, meanwhile, have ABOUTs which are significantly higher than both AWO and AWT; this greater separation between ABOUT and AWT tokens compared to other speakers seems to be due to the lower realizations of the AWT group, which overlaps almost completely with the AWO group for these speakers. (Fig. 5.8). ES also has a significant difference between ABOUT and OUT, as well as a backed ABOUT.



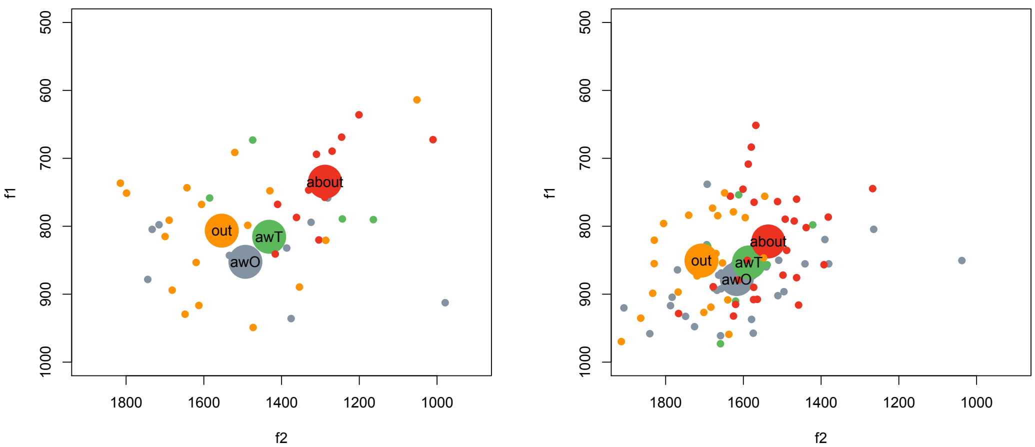
(a) VJ (b) BW

Figure 5.5: Pattern 2: VJ and BW raise *about* and *out*, but not other raising context tokens



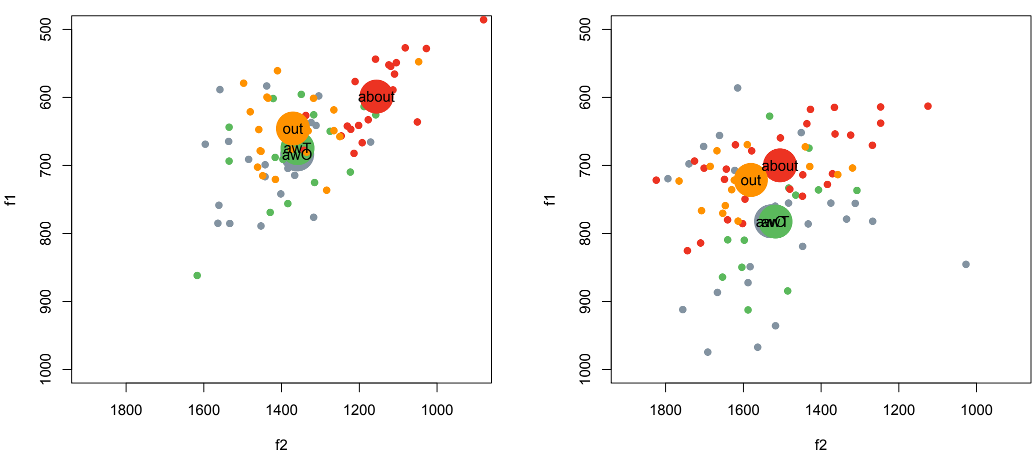
(a) TM (b) LW

Figure 5.6: Pattern 2: TM and LW raise *about* and *out*; *about* is also significantly higher than other raising items, and is backed compared to the other groups.



(a) NW (b) CW

Figure 5.7: Pattern 2: NW and CW raise *about*.



(a) ES (b) JF

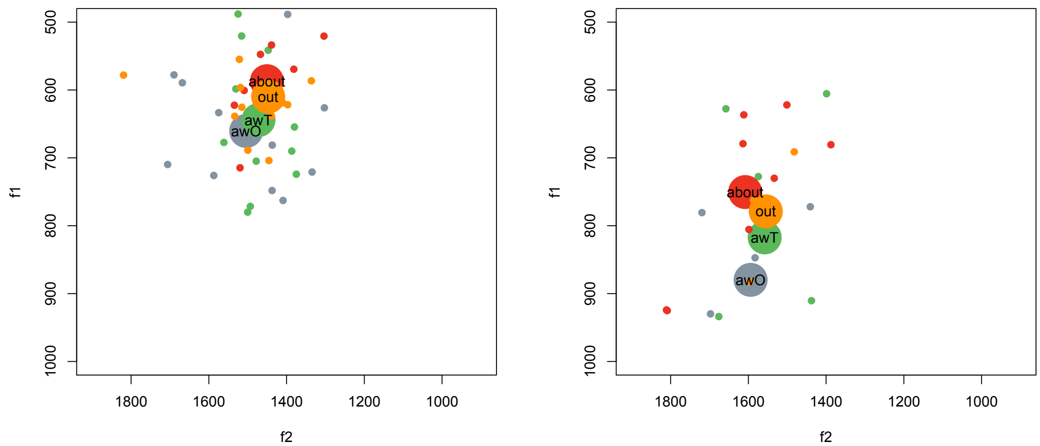
Figure 5.8: Pattern 2: ES and JF raise *about*, which is also significantly higher than other raising items.

5.3.3 Speakers with no significant differences across lexical context

Finally, the analyses for two speakers, JC and BK, showed no significant differences between any of the lexical contexts. For JC, the effect sizes associated with each Raising context are: AWT: -16; OUT: -51; ABOUT:-73. For BK, they are: AWT: -63; OUT: -101; ABOUT: -129. Though some of these effects are quite large, none are statistically significant; in BK's case, this lack of significance may be due to a low token count. However, the placement of the different lexical contexts in phonetic space for each of these speakers is consistent with the patterns already seen in other speakers: ABOUT is most raised, followed by OUT, and then (AW)T, with (AW)O realized lowest in the space.

5.3.4 Summary of results

While there is much fine-grained variation in the results presented above, several generalizations emerge. The majority of speakers show some evidence of Canadian Raising: 7 speakers raise (aw) in all raising contexts, while 8 raise only in the word ABOUT (and sometimes OUT). Across speakers, these contexts generally order themselves in the height dimension in a consistent way: ABOUT is consistently realized highest in the vowel space, followed by OUT, AWT, and then AWO. These patterns are apparent even in the speech of speakers JC and BK, who showed no significant differences between groups.



(a) JC

(b) BK

Figure 5.9: No significant raising: JC and BK show no significant differences between any of the 4 lexical contexts.

5.3.5 The specialness of *out* and *about*

The fact that native Canadians exhibit Canadian Raising in their conversational speech is not entirely surprising. What *is* striking is the behavior of *out* and *about*, which tend to be realized higher in the vowel space than other raising context words, and for nearly half of the speakers, seem to be the only words which show robust raising. These words occur with extremely high frequency in this corpus, accounting for over half of all (aw) tokens collected. According to a usage-based account of change, such high frequency items should be the first to show signs of accommodation towards the new dialect: NYAE tokens of *out* and *about* produced with lowered diphthongs should be accruing more quickly in each speaker's parametric space than lowered tokens of, e.g. *gout* and *couch*, implying that *out* and *about* ought to be among the lowest (aw)T words. Contrary to this prediction, these items resist lowering for nearly all speakers, even when other (aw)T items have lowered.

5.3.5.1 Ruling out duration differences as a possible cause

It is possible, however, that another factor related to frequency can account for the unexpected patterning of *out* and *about*. As noted above, Canadian Raising is a process that occurs in a shortening environment. Vowels occurring before voiceless obstruents in English are shorter than vowels which occur before voiced sounds. Raising in shortened vowels is thus a phonetically natural process: the tongue has less time to fully lower for the nucleus of the diphthong before rising for the offglide, with the result that the realization of the nucleus will be somewhat higher. Of course, Canadian Raising is not a completely phonetic process, as evidenced by the fact that not all dialects/speakers exhibit it. However, there will likely be some small degree of phonetic raising in pre-voiceless contexts for any speaker of English.

Highly frequent items tend to be realized with a shorter duration compared to rare items(e.g. Jurafsky et al. 1998, Gregory et al. 1999). If, because of their higher frequency, tokens of *about* and *out* are often realized with shorter durations compared to other raising context words, then *about* and *out* would also more strongly favor a certain amount of phonetic raising compared to other (aw)T words. Across the 1,210 token corpus, there are small differences in mean duration between the three raising context groups: AWT words have a mean duration of 158ms, while ABOUT and OUT average 154ms and 167ms, respectively; according to a set of pairwise t tests with holm correction, the only significant difference here is between OUT and ABOUT (p=.032). A comparison of the distributions of durations

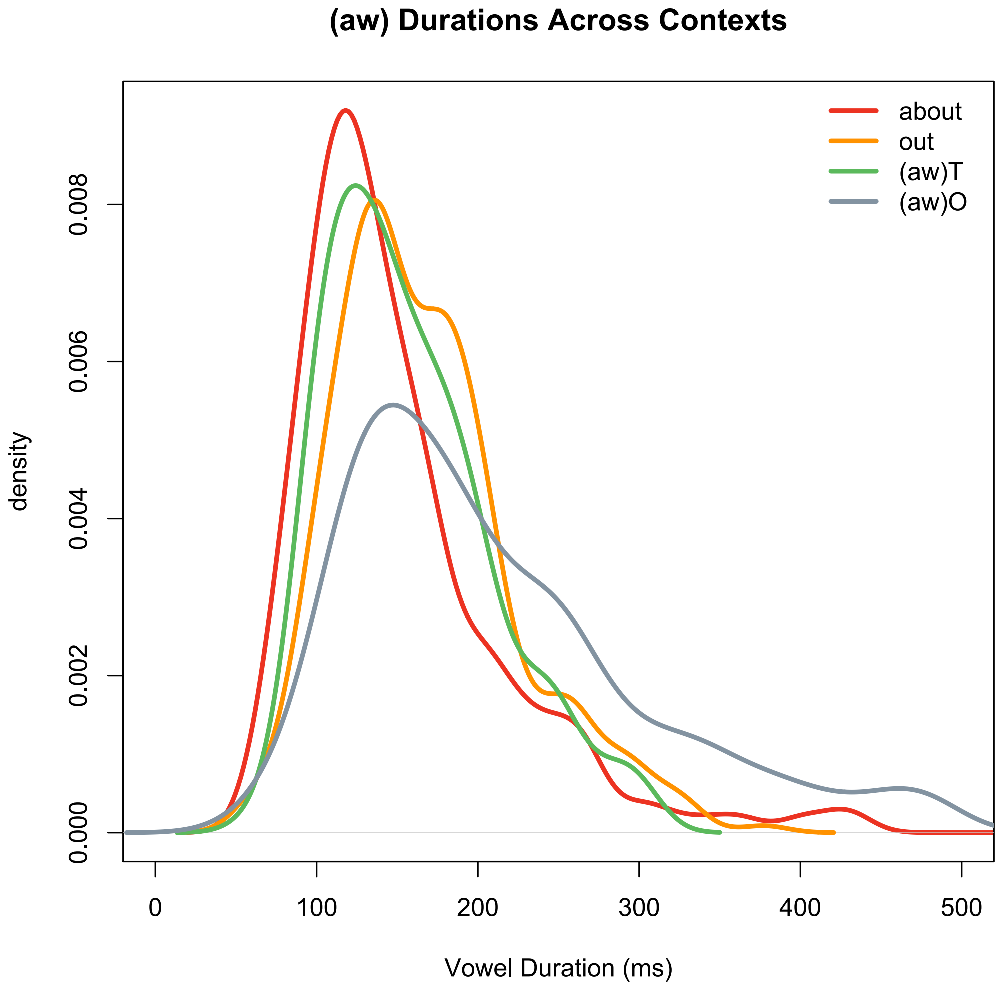


Figure 5.10: Distributions of (aw) durations in the 4 lexical contexts

in ABOUT, OUT, and the AWT items across speakers shows that tokens of *about* tend to be realized with shorter durations compared to the other raising context words (Fig. 5.10).

Is it the case that the excessive raising of *about* compared to the other raising contexts words in this data is due entirely to this durational difference? That is, across raising context tokens - *about*, *out*, and (aw)T words - is duration the sole predictor of F1 values? Or, does *about*ness or *out*ness significantly predict a lowered F1, beyond what duration accounts for? To answer this question for each speaker, a regression model was fit using lm() in R, with F1 as the dependent variable, Duration as the first term in the model, and Lexical Context as the second term in the model. For these analyses, only three lexical contexts were compared: ABOUT, OUT, and AWT. As discussed in Chapter 4, putting Lexical Context second in the model means that this factor will only show significance if it still has predictive value after Duration has already been taken into account. Treatment contrasts were used, with AWT set as the default level. The resulting model shows whether ABOUT and OUT each significantly differ from AWT, after duration has been controlled for.

One general result of these analyses is that Duration is indeed a significant predictor of F1 values for each and every speaker: longer duration is associated with lower vowels, as expected. For each speaker, there are two additional results: whether ABOUT is significantly higher than AWT once this duration effect has been taken into account, and whether OUT is significantly higher than AWT once duration has been taken into account. These results must be considered in light of the results found in the previous section: the issue is whether speakers who showed a significant difference **between about and awT** in the analysis above continue to show this difference once duration has been factored in. A summary of the results is in Table 5.4.

Three of the Pattern 1 speakers (LC, DB, and PW) showed no significant differences between the three raising groups in the previous analysis. As noted in Table 5.4, this continues to be the case once duration has been taken into account. The remainder of the Pattern 1 speakers (GH, LG, SS, and EW) did show a significant difference between ABOUT and AWT tokens in the first analysis, and this result is corroborated here: even with DURATION in the model, ABOUT is still different from AWT. None of these speakers showed a significant difference between OUT and AWT in the first analysis, so it is not surprising that no difference between these groups is detected here.

The Pattern 2 results from the first analysis are similarly supported. NW and CW's ABOUT, though different from AWO words in the first analysis, was not found to be significantly different from AWT, and this is also the case in this pass. Speakers TM, LW, ES, and JF had ABOUT significantly different from AWT words in the first analysis, and this is also the case here; in addition, TM now shows a difference between OUT and AWT. Similarly, the results for BW and VJ differ, though in a way which only affirms the special status of the salient words: while ABOUT and OUT did not differ significantly from AWT in the first analysis, factoring duration into the model here has resulted in ABOUT becoming a predictor of F1 for both of these speakers, and OUT for VJ.

It seems, then, that while duration does affect F1 in this data, duration differences alone do not account for all of the variation in this measure; being ABOUT (and occasionally OUT) also predicts lower F1 values, and therefore higher diphthong nuclei.

	OUT different from AWT Words		ABOUT different from AWT Words	
Speaker	First Analysis	Duration controlled	First Analysis	Duration controlled
BK	no	no	no	no
JC	no	no	no	no
LC	no	no	no	no
DB	no	no	no	no
PW	no	no	no	no
GH	no	no	yes	yes
LG	no	no	yes	yes
EW	no	no	yes	yes
SS	no	no	yes	yes
NW	no	no	no	no
CW	no	no	no	no
LW	no	no	yes	yes
ES	no	no	yes	yes
JF	no	no	yes	yes
TM	**no**	**yes**	yes	yes
BW	no	no	**no**	**yes**
VJ	**no**	**yes**	**no**	**yes**

Table 5.4: Comparison of OUT and ABOUT with AWT words. This table summarizes the significant differences found in F1 between OUT and ABOUT and the AWT words. "First Analysis" refers to the results of the post hoc t tests described in section 5.3, while the columns labelled "Duration controlled" contain the results of the regressions which include Duration as a predictor.

5.3.5.2 More on *out* and *about*

Because a purely durational account of these differences has been ruled out, the unexpected behavior of *out* and *about* must be explained some other way. As noted above, the especially high realization of these highly frequent words runs counter to the frequency predictions made by usage-based phonology, which are that high frequency words should be the most advanced with respect to a change. *out* and *about* should be the first words to lower given exposure to the low (aw)T of NYaE, but instead remain raised.

It is possible that these patterns are the result of entrenchment effects: while the speakers in this study have presumably been been deluged with nonraised (aw) tokens of these words while living among speakers of NYaE, their years living in Canada have also left them laden with many raised tokens of these highly frequent words. If D1 exemplars of these words sufficiently outnumber their D2 counterparts, then this could account for the lack of lowering in these highly frequent words. However, such entrenchment effects were not apparent in the low back vowel data examined in the previous chapter, insofar as more frequent items were in fact shown to to more advanced in accommodation to a NYaE contrast.

As discussed at the beginning of this chapter, *out* and *about* may behave strangely for extralinguistic reasons. To assess the extent to which this is true, it is first necessary to ensure that these items do, in fact, behave atypically, by looking more closely at the patterning of less salient words.

5.3.6 Frequency effects among nonsalient words

The problem posed by *out* and *about* is that these very high frequency words do not behave the way high frequency words are predicted to by a usage-based model. This may be attributable to the saliency of these words, which are strongly associated with the feature of Canadian Raising. Is it the case that other (aw)T words pattern in a more expected way?

To answer this question, an analysis of frequency was carried out on a subset of the (aw) data which contained only the 372 AWO and 208 AWT tokens. As with the low back vowel data, words were coded for Frequency based on a speaker-internal raw count of how many times a lexical item appeared in a specific interview. Data from all 17 speakers was again combined, and lmer() used to fit a series of mixed effects models that contained a random effect of Speaker, in order to correct for gross differences in formant magnitudes between speakers.

A step-up model comparison approach was used to determine whether adding predictors of interest improves the model. The null model *m0* included only a random effect of Speaker as a 'predictor' of F1. The next model, *m1*, added a fixed effect of Lexical Context. Unsurprisingly, a comparison of these two models shows a highly significant effect of adding Lexical Context to the model ($p<2.2e-16$), with AWT tokens being on average about 66Hz higher in the vowel space than AWO tokens.

Next, a main effect of FREQUENCY was added to create *m2*. A comparison of this model with *m1* showed that adding a main effect of FREQUENCY does not significantly improve the model ($p= 0.10$). This is unsurprising, as no overall frequency effect across (aw) tokens is expected: (aw)O words are already low in CE, so are not expected to shift very greatly as a result of exposure to similarly low NYAE (aw)O tokens.

(aw)T, on the other hand, is expected to shift, as low exemplars of NYAE (aw)T are added to the representations of these words: high frequency words

	Effect on F1 (Hz/Count)
awT	-97
Frequency	-1.5
awT*Frequency	8

Table 5.5: The effect of Frequency on F1 in awT words

will accrue new nonraised exemplars at faster rate, resulting in lower productions of these words. An interaction between Frequency and Lexical Context is thus predicted: there should be no effect of Frequency on F1 for (aw)O words, but a negative effective of Frequency of (aw)T, such that high frequency (aw)Ts are realized lower in the vowel space.

To test whether this is the case, a model *m3* was created which included an interaction term Frequency*LexicalContext in addition to the terms of *m2*. A comparison of *m3* and *m2* showed that *m3* is indeed a significantly better model (p=0.002). The effect sizes associate with each term in *m3* are listed in Table 5.5. In this model, the effect associated with being an awT context item as opposed to a awO item is -97Hz; that is, (aw)T items are significantly higher than (aw)O items. The main effect of Frequency on F1 is -1.5052; that is, for items in any context, higher frequency is associated with slight raising (however, given that this main effect did not significantly improve *m2*, it is unlikely to exert a significant effect here). Finally, the interaction term effect indicates that higher frequency (aw)T items are, in fact, realized lower in the vowel space than lower frequency (aw)T items: a one-step increase in Frequency is associated with an 8Hz lowering in awT tokens.

In summary, while the pooled data show that awT items are on a whole still significantly raised compared to awO context words, the degree of raising within the first group is partially predicted by Frequency. Higher frequency raising context items are lower than low frequency items, consistent with the predictions of a usage-based account in which these items are undergoing lexically gradual shift towards lower NYaE-like realizations.

5.4 Discussion of results

This chapter examined the realization of (aw) in pre-voiceless contexts in the speech of 17 Canadians living in the New York region. While Canadian Raising is perhaps the most salient linguistic feature that characterizes CE, it is not a feature of NYaE. The analysis presented here sought to determine whether these speakers have accommodated to (aw) as produced in the new dialect - that is, whether they have ceased or at least attenuated their use of Canadian Raising - and whether there is any evidence that this accommodation is happening on a lexically gradual basis.

As with the low back vowels, the answer to the question of whether accommodation has occurred is a qualified yes. Some speakers indeed seem to have curtailed their use of raised (aw) in (aw)T words. However, this feature also shows remarkable stability: several speakers continue to show raising in all contexts, and essentially everyone continues to raise the diphthongs in *out* and *about*. Moreover, the difference between raising and nonraising contexts is not only statistically significant in these cases, but of appreciable magnitude, often well above the 60Hz threshold used by Labov et al. (2006) to categorize speakers as exhibiting Canadian Raising.

Despite this stability, however, there is evidence of shift towards the lowered realizations of NYaE. Across speakers, higher-frequency raised items (leaving aside *out* and *about*) are realized with lower nuclei. These frequency effects, and the apparently gradual nature of the shift towards NYAE norms that they indicate, are consistent with a usage-based model in which speakers incorporate their exposure to the new dialect in formulating targets for their own productions. A generative account of these results would require lexically specific quantifications of the phonetic values of the Canadian Raising rule in order to model these results – a modification which would make it notationally very similar to Exemplar Theory.

The exception to this overall pattern of shift is the extremely frequent and extremely salient word *about* and, to a lesser extent, *out*, which resist lowering: all speakers produce these words somewhat higher than other raising context words, regardless of their overall raising pattern. The exceptional behavior of these items, and an explanation for their patterning, will be further discussed in Chapter 8.

CHAPTER 6

LOW BACK VOWELS VS. CANADIAN RAISING

6.0 Introduction

The previous two chapters described how the 17 speakers in this study behave with respect to two phonological features: low back vowel contrast and Canadian Raising. Chapter 4 described how several speakers show evidence of having acquired a small low back vowel distinction, producing a significant phonetic difference on one or more dimensions that is not predictable by phonological context alone; across speakers, this change seems to be proceeding in a lexically gradual way. Chapter 5 showed that most speakers exhibit (aw)-raising in some form, but overall show frequency effects indicating that raising context words are gradually shifting towards lower realizations. For both features, the native dialect is still clearly influential: all speakers maintain a low back merger in perception and production in minimal pair contexts, and continue to exhibit raising in at least some words. However, in both cases there is evidence of a lexically and phonetically gradual accommodation to NYaE realizations.

Thus far, the results for each feature have been considered separately. However, these two sets of data do not represent 34 independent results. Instead, they are paired by speaker: SS's low back vowel results and SS's raising patterns are components of the same phonological system, and have developed in the same psychological and social milieu. It is reasonable to ask, then, whether there is any relationship between these two features, and if so, how this relationship may be accounted for within a phonological theory of intraspeaker change.

6.1 Recap of predictions

Chapter 3 described two types of phonological theories - generative phonology and usage-based phonology - and set out the predictions that each type of theory makes regarding how speakers should go about accommodating to the two dialect features which are the focus of this study. Most of these predictions pertained either to (o)/(oh) or raising alone. However, the two theories also make different predictions regarding the relative acquirability of these features.

In generative theory, accommodation to each of these features happens at different levels, via different mechanisms: true acquisition of a contrast requires many independent changes to underlying lexical representations, while altering a phonological rule involves potentially only one change to the rule itself. This predicts an asymmetry in accommodation: altering or losing the raising rule should be easier than, and probably precede, noticeable acquisition of the low back vowel contrast. Looking across speakers as a group, more would be expected to have lost or attenuated their raising rule than to have acquired a low back vowel contrast; a stronger prediction would be that acquisition of an (o)/(oh) contrast implies that loss of raising has occurred (though not vice versa).

In usage-based theory, both types of changes happen at the same level, via the same mechanism: gradual accrual of new exemplars of the relevant word forms. In this theory, no asymmetry is predicted: accommodation to both types of features should happen at more or less the same rate. Looking across speakers, we would expect accommodation towards these features to be positively correlated: success at gaining a low back vowel contrast would be associated with success in losing the raising rule.

The next two sections will test these predictions, by examining the relationship between (o)/(oh) and raising from both qualitative and quantitative perspectives.

6.2 A qualitative comparison

I will start by considering the speaker data from a qualitative perspective: who has acquired a low back vowel contrast, who has suppressed Canadian Raising, and does a particular pattern of behavior with respect to one of these features at all predict behavior regarding the other?

Considering a binary categorization of just these two features, there are 4 logically possible systems (shown in Table 6.1). A speaker might exhibit Canadian Raising and not distinguish (o)/(oh), reflecting a system that is still essentially CE. Or, she may produce (aw) with no raising before voiceless segments and contrast (o)/(oh), reflecting a fully successful[1] accommodation to NYAE. More interestingly, the speaker may exhibit one of two mixed systems, using the CE setting for one feature and the NYAE setting for the other. These mixed dialects might be transitional states in the move towards the lower righthand quadrant of Table 6.1, or they might be endpoints in themselves.

[1]in this simple, two-dimensional system!

	(o)/(oh) same	(o)/(oh) different
(aw)T raised	**Canadian English**	mixed/transitional
(aw)T lowered	mixed/transitional	**New York-area English**

Table 6.1: Logically possible types of systems

	(o)/(oh) same	(o)/(oh) different
(aw)T raised	✓	no speakers
(aw)T lowered	most speakers	✓

Table 6.2: Accommodation Patterns Predicted by a Generative Account

The previous section briefly summarized how generative and usage-based theories predict different patterns of accommodation towards these two features. These predictions can be restated in terms of how speakers are expected to be distributed among the 4 system types described in Table 6.1. In both accounts, there will be speakers who fit squarely in the CE or NYaE English cells: some speakers may not have had enough input to make progress towards accommodation in either feature, and some speakers may have had enough input to be able to successfully adopt both. The existence of speakers of either of these types will not by itself be sufficient for deciding between these two theories.

What will be informative is the distribution of speakers among the remaining two quadrants. According to a generative account, the majority of, if not all, transitional speakers should exhibit a system in which (aw)T has been lowered, but (o) and (oh) are still merged, reflecting the relative difficulty of acquiring these two features as described above. There should be few, if any, speakers showing the opposite system, which maintains raising while also showing acquisition of a low back contrast. This predicted distribution is summarized in Table 6.2.

	(o)/(oh) same	(o)/(oh) different
(aw)T raised	✓	x speakers
(aw)T lowered	x speakers	✓

Table 6.3: Accommodation Patterns Predicted by a Usage-based Account

A usage-based account does not imply this sort of asymmetry. Acquisition of a low back contrast and loss of raising should both proceed gradually, and be positively correlated, with neither feature the obvious 'winner' with respect to speed of accommodation. For this reason, speakers are expected to be evenly distributed across the two transitional systems (Table 6.3): depending on the specific tokens of each word class that a particular speaker happens to hear, one feature might temporarily inch ahead of the other in its degree of accommodation, but there is no phonology-internal reason why one feature might reach its tipping point and change its qualitative status before the other.

So, how are the speakers in this study actually distributed across the four possible system types? First, it is necessary determine how speakers should be categorized based on the results from Chapters 4 and 5. With respect to the (o)/(oh) contrast, this is easily done by considering any speaker who showed a significant difference along either (or both) F1 and F2 in conversational context to have these groups "distinct", and speakers with no significant difference along either dimension to have these word classes "the same". By this method of categorization, 11 speakers will fall under the "(o)/(oh) different" column, and the remaining 6 will be under "(o)/(oh) same."

Determining who raises or does not raise is less straightforward. As seen in Chapter 5, most speakers in the group (15 of 17) exhibited raising in at least 1 raising context, but only a subset of these raised in *all* raising contexts. Depending

	(o)/(oh) same	(o)/(oh) different
(aw)T raised	PW	DB, EW, GH, LC, LG, SS
(aw)T lowered	CW, ES, NW, TM, VJ	BK, BW, JC, JF, LW

Table 6.4: Accommodation Patterns Observed: (o)/(oh) vs. (aw)T raising

on where the bar is set for raising, the distribution of speakers across the four systems will look very different.

I will begin by defining "raising" conservatively: a speaker is considered to exhibit Canadian Raising if he shows evidence of raising in all three of the raising groups analyzed in Chapter 5 - that is, if he is a Pattern 1 Unconstrained Raiser. Using this criterion, the speakers are distributed as in Table 6.4. As expected, there are speakers who show evidence of accommodation to both features (BK, BW, JC, JF, LW) as well as a speaker who has accommodated to neither (PW). What is interesting is the distribution across the two mixed/transitional systems: the 11 speakers who show accommodation with respect to only one feature are nearly evenly split between these two cells. On the face of it, this seems to bear out the usage-based predictions outlined above, as there is obvious no asymmetry between these two systems.

The picture changes dramatically, however, when the raising criterion is loosened to include Pattern 2 speakers, who raise in a restricted set of contexts (namely, in *about* and/or *out*). This new criterion, which defines nearly all of the speakers as raisers, unsurprisingly results in a mass exodus of data points from the bottom row of the table. What is striking, however, is the resulting distribution across the transitional cells: the (o)/(oh) distinct, (aw)-raised cell is now the most populous, while the (o)/(oh) same, (aw)-lowered cell has no speakers (Table 6.5). This pattern does not fit the predictions of either a generative or usage-based account.

	(o)/(oh) same	(o)/(oh) different
(aw)T raised	PW, CW, ES, NW, TM, VJ	DB, EW, GH, LC, LG, SS, BW, JF, LW
(aw)T lowered		BK, JC

Table 6.5: Accommodation Patterns Observed: (o)/(oh) vs. raising anywhere

There is an obvious asymmetry between the mixed systems (contrary to the usage-based account), but this asymmetry favors speakers who acquire a contrast *before* losing Canadian Raising (contrary to the generative account).

This result is puzzling, and further corroborated by an examination of the quantitative relationship between these two features.

6.3 A quantitative comparison

While it is possible to categorize speakers in terms of *whether* they have accommodated to each of the new dialect features in this study, Chapters 4 and 5 showed that accommodation to both is a gradient process. It is therefore possible to assess and compare *how much* each feature is present, and see whether there is an expected relationship between these magnitudes.

Again, values for each feature must be assigned to each speaker; in the previous section, this value was binary ('yes/no'), but here a continuous measure that reflects the gradient nature of each of these features will be used. For the (o)/(oh) contrast, a measure of the real distance between (o) and (oh) - that, is, the difference that remains between these groups after phonological factors have been taken into account - is required. The analyses of Chapter 4 returned two such differences for each speaker, in the form of the effect size of Word Class on F1 and F2; these effect sizes can be interpreted as the real difference in height and backness, respec-

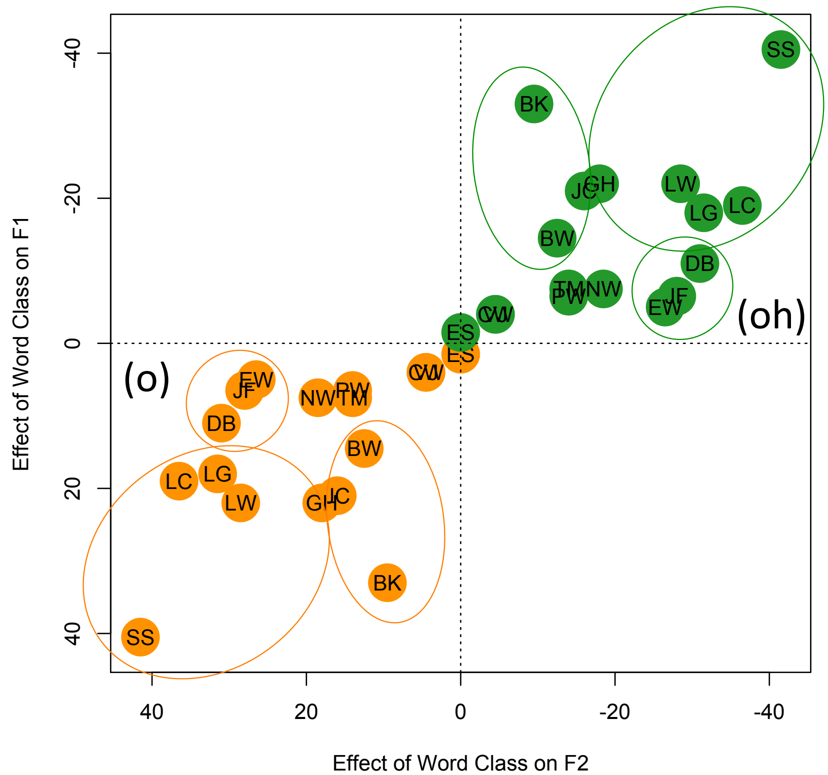


Figure 6.1: Real distance between (o) and (oh) by speaker, based on the effect of Word Class on F1 and F2

tively, in an idealized vowel space. These differences were presented graphically in the Figure 4.23, reproduced in Fig. 6.1. For the single measure of the difference between (o) and (oh), I used the Euclidean distance derived from these F1 and F2 effect sizes.

For raising, the situation is again more complicated. The chapter 5 analyses yielded 3 measures of raising per speaker, in the form of an effect size associated with each raising context group. In each case, the baseline context for comparison was the group of nonraising (awO) context tokens. The effect on F1 associated with each raising context group is thus a measure of how much raising that group shows compared to the (aw)O group. Rather than choose only one measure to compare with low back vowel contrast, I will examine each in turn.

Given these ranges of values corresponding to each feature, with comparable units,[2] it is possible to determine whether these are correlated in any way. Strictly speaking, there does not seem to be any reason to expect a direct correlation between these two features. However, both are expected to have a particular relationship to time. The longer that a speaker is immersed in NYAE, the more exposure she will have to distinct (o)/(oh), so the gap between her low back vowels should widen over time. Similarly, with more time and exposure to NYAE low (aw)T, the distance between raising context tokens and nonraising context tokens should shrink. Removing the temporal middleman, a negative correlation between these two features is expected: as the difference between low back vowels increases, the difference between nonraising and raising groups should decrease.

Surprisingly, this prediction is not borne out. Fig. 6.2 plots the Euclidean effect size of word class against the F1 effect size associated with the AWT group (that is, raising context words which are not *about* or *out*). In fact, there is a significant *positive* correlation between these two features: those speakers who show the greatest distance between (o) and (oh) are also those who show the most raising in nonsalient (aw)-raising words (Spearman's r=.517, p=.036)

This positive relationship is even stronger between (o)/(oh) and raising in the salient words *out* and *about*. For *out*, the correlation with low back distance is .586 (p=.015) (Fig. 6.3), while for *about*, Spearman's rho is .699 (p=.002) (Fig. 6.4).

[2]Both the raising effects and Euclidean distance measures are in Hz.

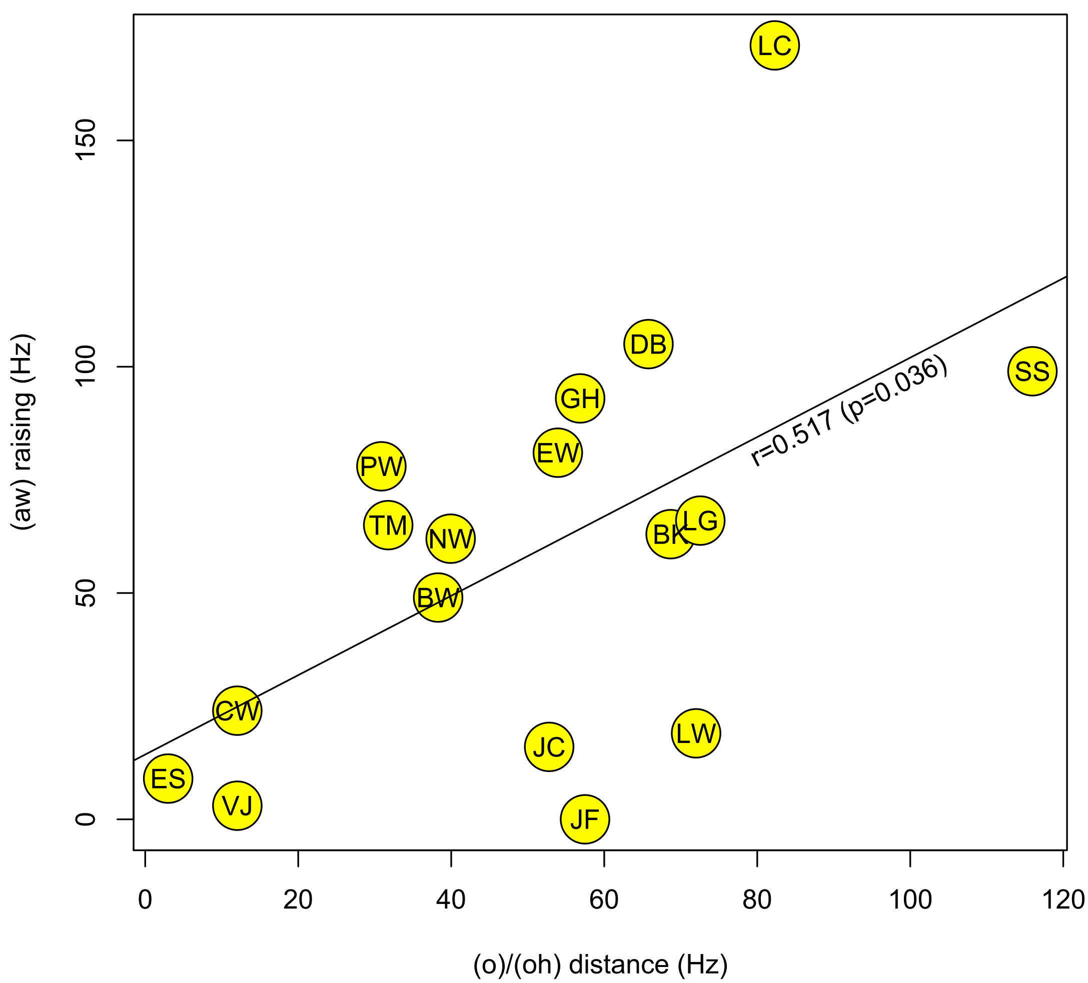


Figure 6.2: (o)/(oh) distance vs. (aw) raising in AWT words

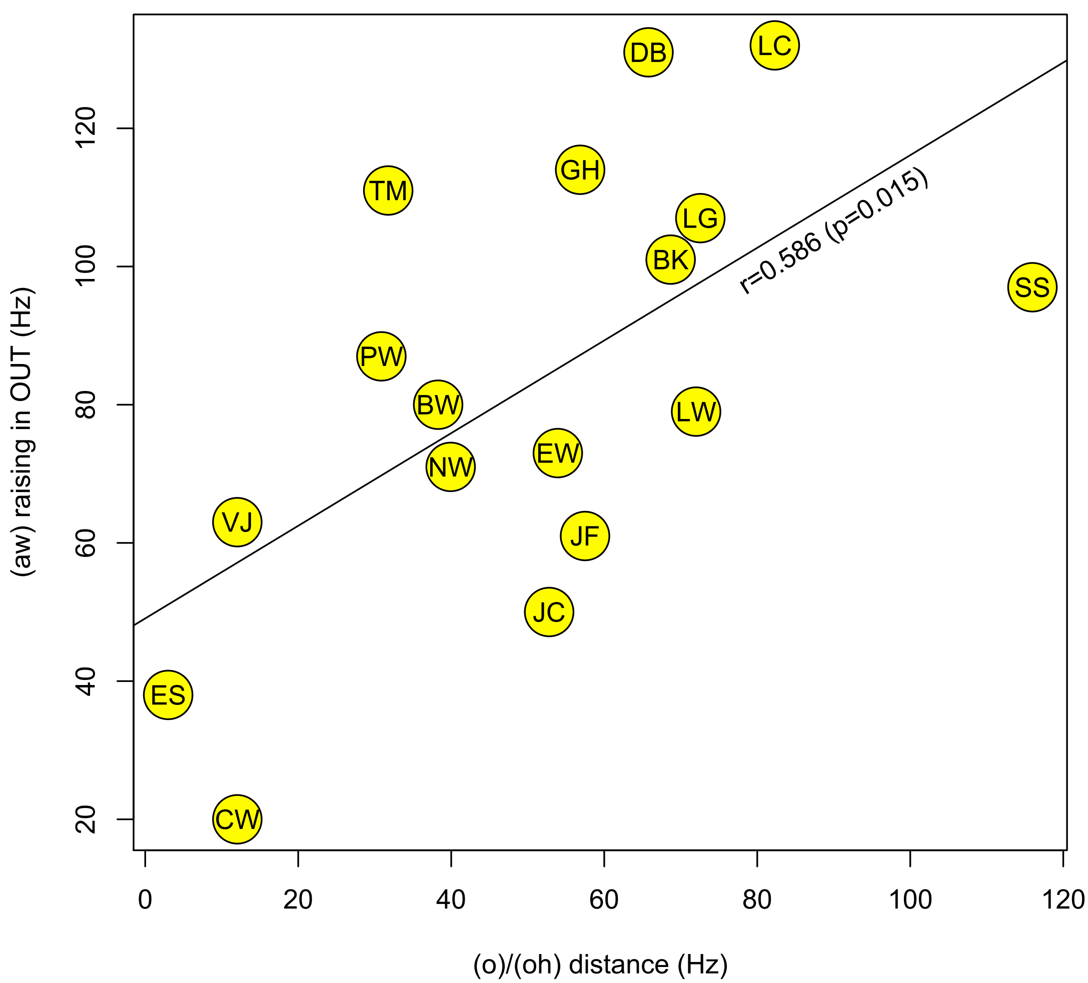


Figure 6.3: (o)/(oh) distance vs. (aw) raising in OUT

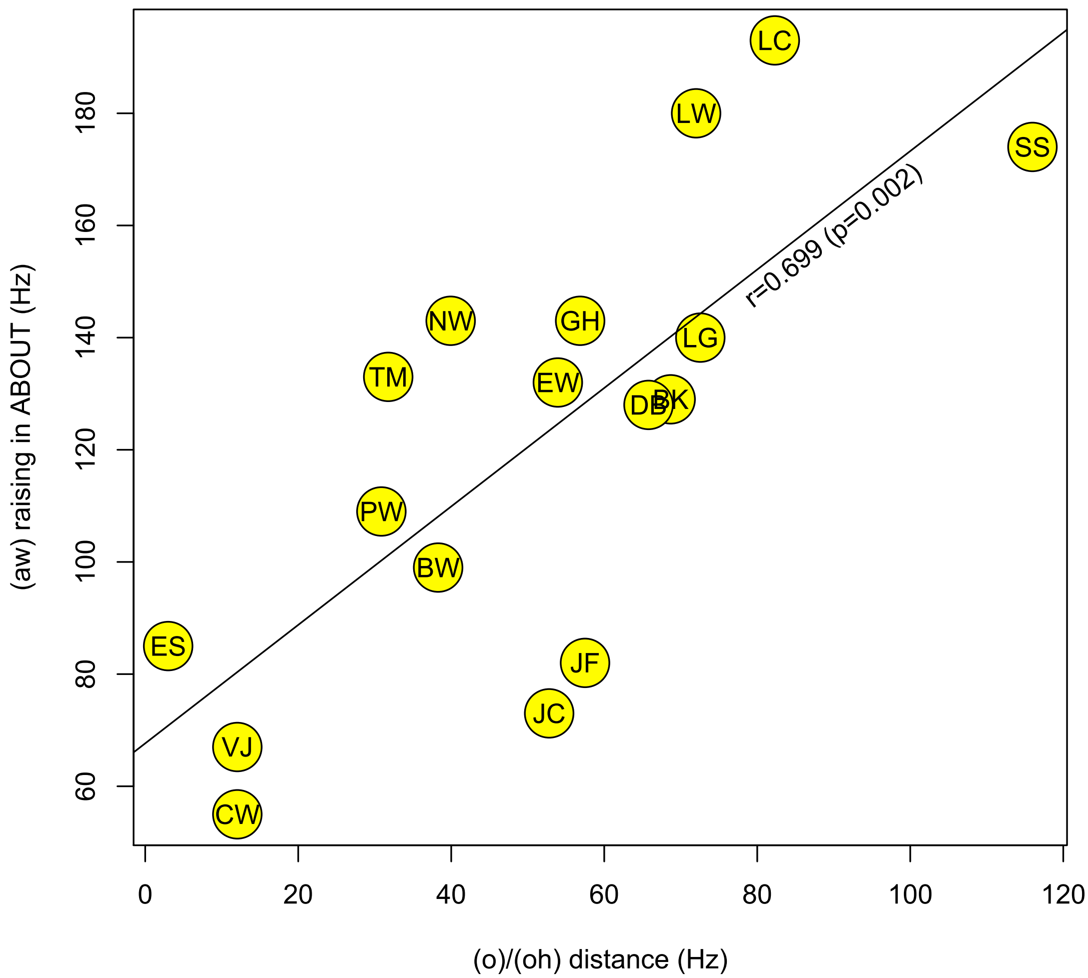


Figure 6.4: (o)/(oh) distance vs. (aw) raising in ABOUT

6.4 Discussion

The analyses in this chapter sought to determine whether there is any relationship between low back vowel distinction and (aw) raising across the group of speaker in this study. Chapter 3 discussed how the two types of phonological theory being tested make different predictions regarding how speakers would be expected to accommodate to these features. The different status of these features in generative theory leads one to expect an asymmetry in accommodation: speakers are expected to have an easier time altering a single phonological rule than changing many underlying representations, so loss of (aw) raising should precede acquisition of a (o)/(oh) contrast. In usage-based theory, however, no such asymmetry is predicted: because loss of raising and acquisition of an (o)/(oh) difference proceed via the same mechanism (gradual shift of relevant word-level representations), speakers are expected to accommodate to both features to more or less the same extent.

A comparison of these features at the qualitative level showed that every logically possible combination of accommodation and non-accommodation is present among these speakers, given a conservative definition of raising: some speakers have both acquired a low back vowel distinction and lost Canadian Raising in most contexts, one speaker fails to show a (o)/(oh) difference and still raises, and the remaining speakers are evenly split between those who have acquired a (o)/(oh) difference but still raise, and those who have stopped raising but still show no low back vowel difference. If speakers who raise in any context are considered to exhibit Canadian Raising, then half of the speakers in the study are Canadian Raisers who also have acquired a (o)/(oh) distinction.

The apparent positive relationship between (o)/(oh) contrast and maintenance of Raising is even more clearly brought out in the comparison of the degree to which each feature is exhibited in each individual's speech. There is a significant positive correlation between degree of low back vowel distinction and degree of Canadian Raising: those speakers who produce the greatest difference between (o) and (oh) also raise pre-voiceless (aw) to the greatest extent. That is, greater accommodation to one *new* dialect feature (low back distinction) is associated with stronger maintenance of an *old* dialect feature (Canadian Raising).

From a purely linguistic perspective, this finding is unexpected. From a usage-based standpoint, low back vowel distance and Canadian Raising ought to be negatively correlated, reflecting a steady influx of new dialect exemplars causing (o) and (oh) to diverge while (aw)T and (aw)O converge. From a generative standpoint, no particular correlation is expected, as these features are not supposed to change gradually in any case; however, we might expect some sort of implicational relationship to hold between these features, with success in acquiring a low back vowel distinction implying that Canadian Raising would already have faded. Neither theory seems to provide a ready account of why speakers are patterning in the way observed here.

A key difference between these features which may be affecting their realizations (and obscuring the phonological predictions) is their differing social meaning. While Canadian Raising is a highly salient marker of Canadian English, low back vowel merger or distinction as such is not strongly associated with any particular dialect or identity. In Chapter 8, I will outline an account of the results presented here that incorporates this social differentiation. But first, Chapter 7 will consider some of the other extralinguistic factors which may contribute to these results.

CHAPTER 7

EXTRALINGUISTIC FACTORS

7.1 Introduction

This study has focused primarily on the linguistic aspects of second dialect acquisition. However, speakers are not merely vessels for sound categories who occasionally bump into each other and share their contents. Social characteristics, experiences, language awareness, and attitude also affect linguistic behavior.

As noted in Chapter 2, the speakers in this study were not selected with the goal of testing the effects of any of these social factors. Indeed, I specifically sought out speakers who were likely to have the experiential and attitudinal characteristics that would favor accommodation to NYaE: those who had expressed a positive attitude towards their new home, who had found a local spouse or best friend, or who had intention to stay in the New York region. Moreover, the group is demographically rather homogenous: nearly all speakers are white, middle class, and have had at least some university-level education.

Even so, the speakers in this study do vary socially along several potentially interesting dimensions. This chapter will investigate the extent to which some of these extralinguistic factors favor or disfavor adoption of NYaE forms.

7.2 Place of origin

While of the speakers in this study acquired their first dialect of English in Canada, they were born and raised in several different regions within this country. As noted in Chapter 2, the variety of English spoken in Canada is considered to be rather homogenous, especially compared to the diverse varieties found throughout the United States. However, recent work by the Phonetics of Canadian English project (PCE, Boberg 2008 has begun to investigate regional variation and change in CE using speech data collected from McGill undergraduates hailing from a variety of areas across Canada. While the PCE does not report any regional variation in the (o)/(oh) merger, there is some indication that Canadian Raising may be present to varying extents across the nation.

This variation is quantitative rather than qualitative: according to Boberg (2008), "the PCE data indicate that Canadian Raising is a largely uniform feature of Canadian English", with 88% of speakers showing an F1 difference of at least 50Hz between (aw)T and (aw)O in word list style. In fact, many speakers exhibited raising to a far greater extent than this, as the mean F1 distance across all speakers was 142Hz.[1] Though this feature is near-ubiquitous across CE, significant regional effects on F1 distance were found. However, Boberg points out that this result is mostly due to the behavior of speakers of Newfoundland. While the 6 speakers

[1]Height and backness of (ay) was also examined, with raising of this diphthong found to be similarly robust among the speakers in the PCE sample.

From	(o)/(oh) same	(o)/(oh) different
Western provinces	ES, NW, PW	EW, JF
Ontario	TM, VJ	BK, BW, LC, LG
Montreal	CW	GH, JC, SS
Eastern provinces		DB, LW

Table 7.1: Speaker region of origin vs. (o)/(oh) status

from Newfoundland had an average F1 distance of only 37 Hz, the mean distance for all other regions ranged from 117-176Hz. Amongst the non-Newfoundland regions, only the difference between Quebec (117Hz) and the Prairies (168Hz) was found to be (marginally) significant.

As none of the speakers in this study acquired CE in Newfoundland, I have avoided a potentially major source of regionally-based variation in both (aw)-raising and (o)/(oh). However, because there may be slight variation in the degree of raising across other regions of Canada, it is worth checking whether region of origin may account for some of the interspeaker variation observed here.

The majority of speakers in this study come from either Ontario (often, Toronto or nearby areas) or Montreal[2]. The remaing speakers moved from a wider variety of regions; in order to examine more general regional patterns, these speakers were categorized as coming from either Western provinces (British Columbia, Alberta, Saskatchewan, Manitoba), or Eastern provinces (New Brunswick, Nova Scotia). Tables 7.1 and 7.2 the distribution of speakers across these four regional groups in terms of their qualitative behavior with respect to low back vowel contrast and (aw)T raising.

[2]These are the two largest population centers in Eastern Canada, so an abundance of speakers from these areas could hardly be avoided.

From	raising	no raising
Western provinces	ES, PW	EW, JF, NW
Ontario	LC, LG	BK, BW, TM, VJ
Montreal	GH, SS	CW, JC
Eastern provinces	DB	LW

Table 7.2: Speaker region of origin vs. presence of (aw)T-raising

Based on the distribution of speakers in Table 7.2, there does not seem to be any relationship between region of origin and presence of (aw)T raising. However, there is the slightest suggestion of a West to East increase in low back vowel contrast (Table 7.1); while speakers from the Western provinces seem to slightly favor a merged (o)/(oh), most speakers from Ontario and Montreal show an (o)/(oh) difference, and both Eastern speakers make this contrast. This relationship is not, however, statistically significant (based on a Fisher's Exact Test), which is unsurprising given the small sample size.

An examination of place of origin and the quantitative realization of these features also reveals no apparent regional variation with respect to (aw)T-raising, yet a possible relationship between region and difference in (o)/(oh). Fig. 7.1 reproduces the scatterplot of effect sizes from Chapter 6, color coding each speaker according to place of origin. While the Montreal and Ontario speakers are distributed fairly evenly through this two-dimensional space, all 5 speakers from Western provinces have less than 60 Hz distance between (o) and (oh), while both Eastern speakers have more than 60 Hz distance between these vowels. This patterning could reflect the existence of pre-existing regional variation in the low back vowel system within Canada, contrary to the dialectological consensus that the (o)/(oh) merger is well-established and stable in CE.[3]

[3]This pattern seems to mirror the West to East difference in low back vowel realization in the United States. If such a West-East difference were to be found in Canada, it would indicate that

From	(o)/(oh) same	(o)/(oh) different
Female	CW, NW, TM, VJ	BK, DB, JF, LC, LG, LW, SS
Male	ES, PW	EW, GH, JC

Table 7.3: Speaker gender vs. (o)/(oh) status

However, in this case it is difficult to tease apart the effect of place from at least one other potentially relevant social factors: specifically, both of the (o)/(oh)-contrasting Eastern speakers have spouses who are from (o)/(oh) contrasting regions of the United States., while none of the Western province speakers have partners with a 2P system (see Section 7.5).

Whether the slight regional difference suggested here indicates the existence of a real place effect or is merely a coincidence that will disintegrate with more data, these patterns indicate the need for further investigation of the effect of region of origin.

7.3 Gender

The 17 speakers in this study are not balanced by gender: only 6 are male. However, because gender is often an important factor in linguistic variation and change, it is worth seeing whether acquisition of the NYAE features of interest shows any gender patterning. As shown in Tables 7.3 and 7.4, there is no obvious relationship between gender and qualitative behavior with respect to either variable.

The quantitative patterns are more suggestive. Fig. 7.2 again plots each speaker according to (o)/(oh) distance and degree of (aw)T-raising, with speak-

the U.S.-Canadian border has become permeable to certain phonetic features which had been previously been blocked from diffusion (Boberg 2000).

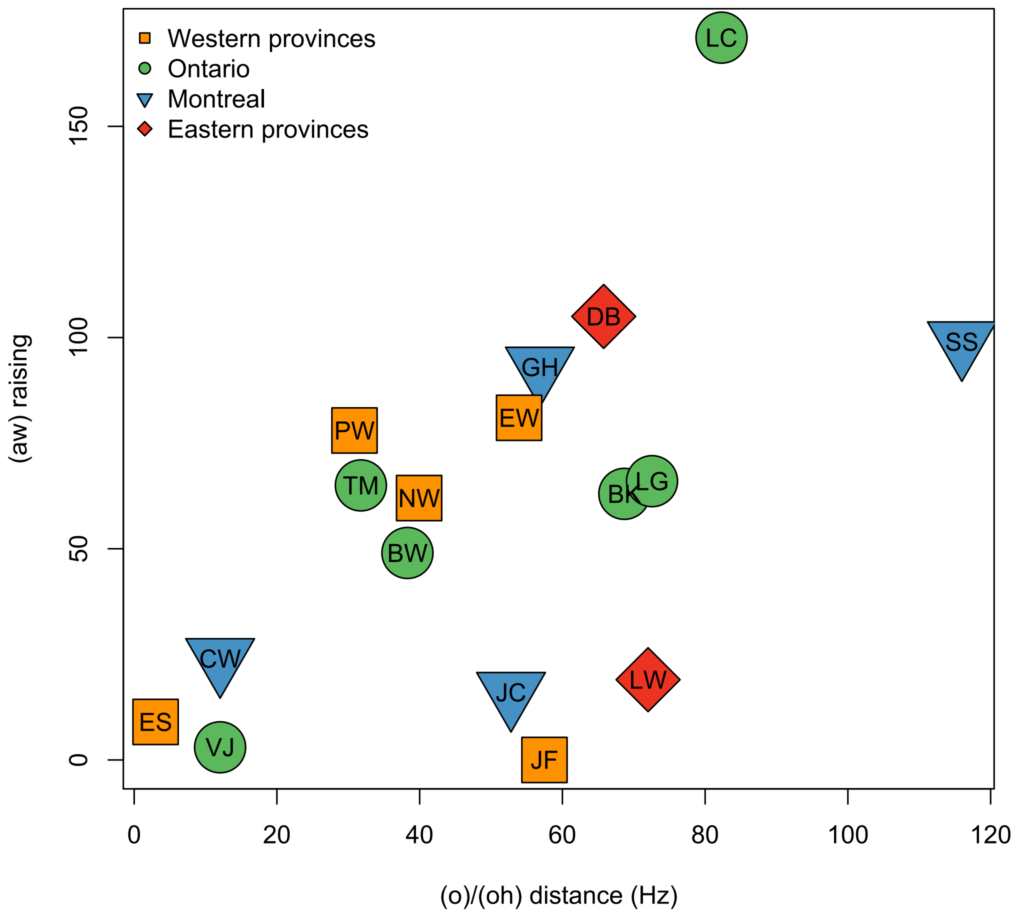


Figure 7.1: (o)/(oh) distance vs. (aw)T raising, by region of origin

From	raising	no raising
Female	DB, LC, LG, SS	BK, CW, JF, LW, NW, TM, VJ
Male	EW, GH, PW	ES, JC

Table 7.4: Speaker gender vs. (aw)T raising status

ers color-coded by gender. While females are distributed throughout the two-dimensional space, men are more constrained, showing less than 60Hz difference between (o)/(oh) and less than 100Hz raising in nonsalient raising contexts. This may reflect a real gender difference, such that men are (for whatever reason) more prone to attenuate Canadian Raising as well as less likely to adopt a strong (o)/(oh) contrast. Again, however, this factor is inextricably linked to another possibly relevant social characteristic: none of the men in this study have a partner with a low back vowel contrast (see Section 7.5).

7.4 Time spent in the New York region

One key factor that has been observed to affect both second dialect acquisition and second language acquisition is time spent in the new region. More time spent among speakers of the new variety means more exposure to that variety, which increases the likelihood that new variants will be acquired (e.g. Payne 1976).

The speakers in this study vary in terms of how long they have lived in the New York region. The newest arrival is PW, who had only been in New York for about 3 months at the time of interview; the oldest is VJ, who had been living in Manhattan for over 40 years. Though the remaining speakers are not distributed evenly between these extremes, there is enough variation to investigate the relationship between time spent in the new region and acquisition of each of the NYaE features discussed here.

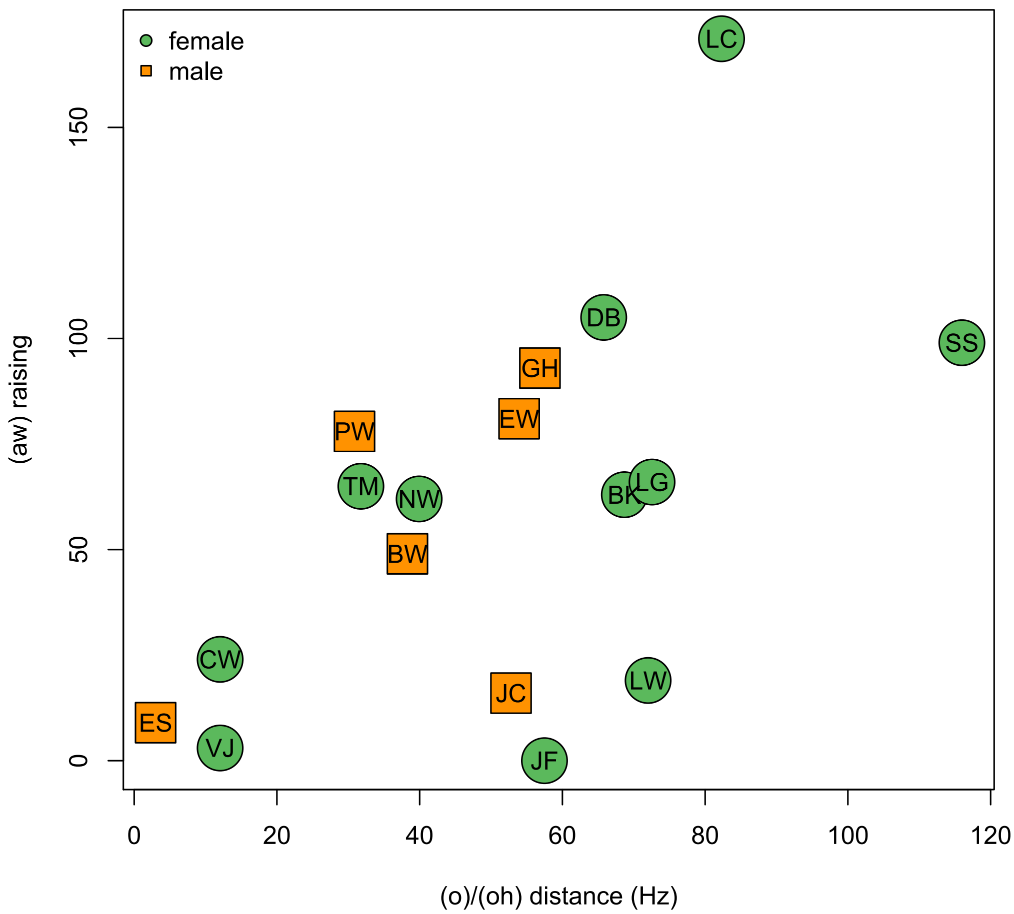


Figure 7.2: (o)/(oh) distance vs. (aw)T raising, by gender

7.4.1 Time and (o)/(oh) distance

The obvious first relationship to examine is that of the raw number of years spent in the NY/NJ region and realization of a local feature. Figure 7.3 plots the Euclidean distance between (o) and (oh) for each speaker against the number of years that speaker had been living in the New York region at time of interview. There is no visible relationship between these two variables, and no statistically significant one; Spearman's rho is a meager -0.007 (p=0.981).

It may be the case, however, that this measure of exposure is too simple. While more years in a new region may result in more input from the associated new dialect, it is also important to take into account years spent in the old dialect region: extensive exposure to the old dialect could favor entrenchment, resulting in slower accommodation to the new variety. Fig. 7.4 visualizes the relationship between old dialect exposure and (o)/(oh) by plotting Age Moved to the New York region (that is, number of years spent in Canada) against low back vowel distance.

While Age of Move has been shown to be a significant factor in studies of child second dialect acquisition, e.g. Payne (1976), it would be surprising if this factor turned out to be significant for this group of speakers, who all moved to the New York region after the age of 21. Indeed, as Fig. 7.4 shows, this factor has no effect (r=0.02, p=0.944).

A more sophisticated measure would incorporate both years of exposure to the new dialect and years of exposure to the old; to capture both of these in a single value, I calculated the proportion of life spent in the new dialect region by simply dividing the number of years each speaker has spent in New York by the speaker's age.. Fig. 7.5 plots this proportion against low back vowel distance. However,

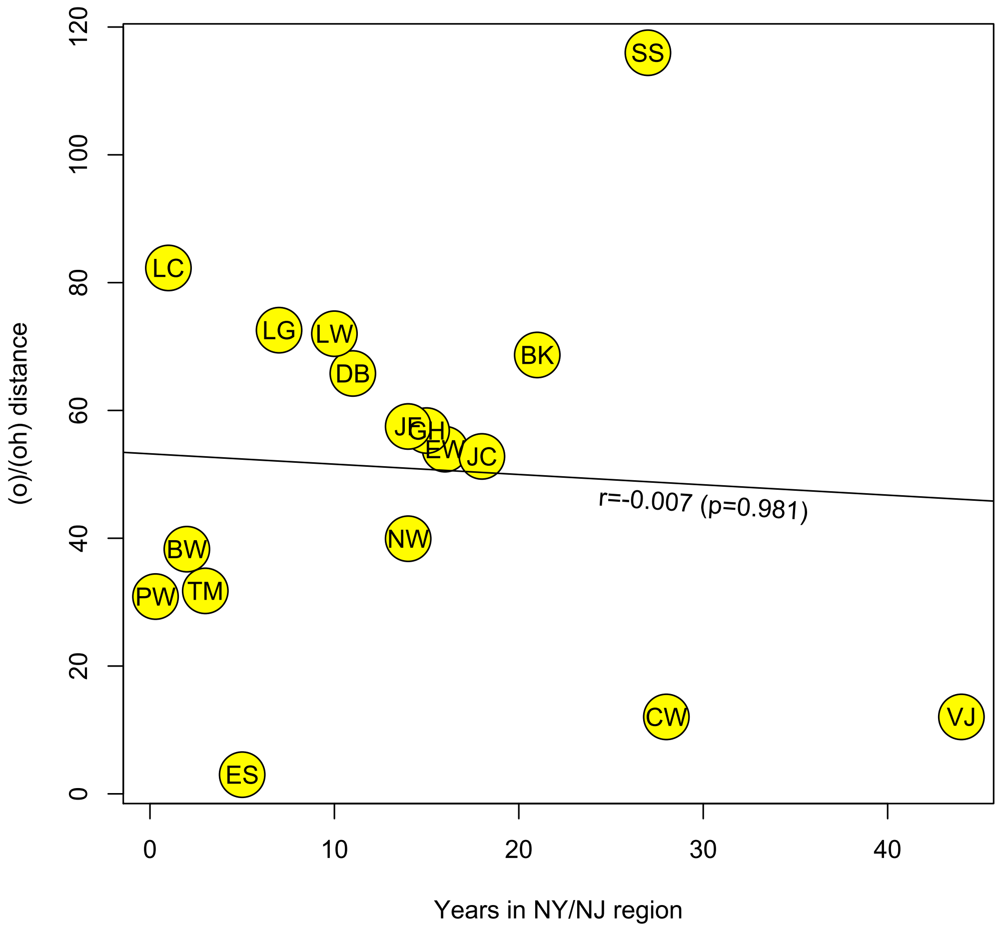


Figure 7.3: (o)/(oh) distance vs. years in NY/NJ

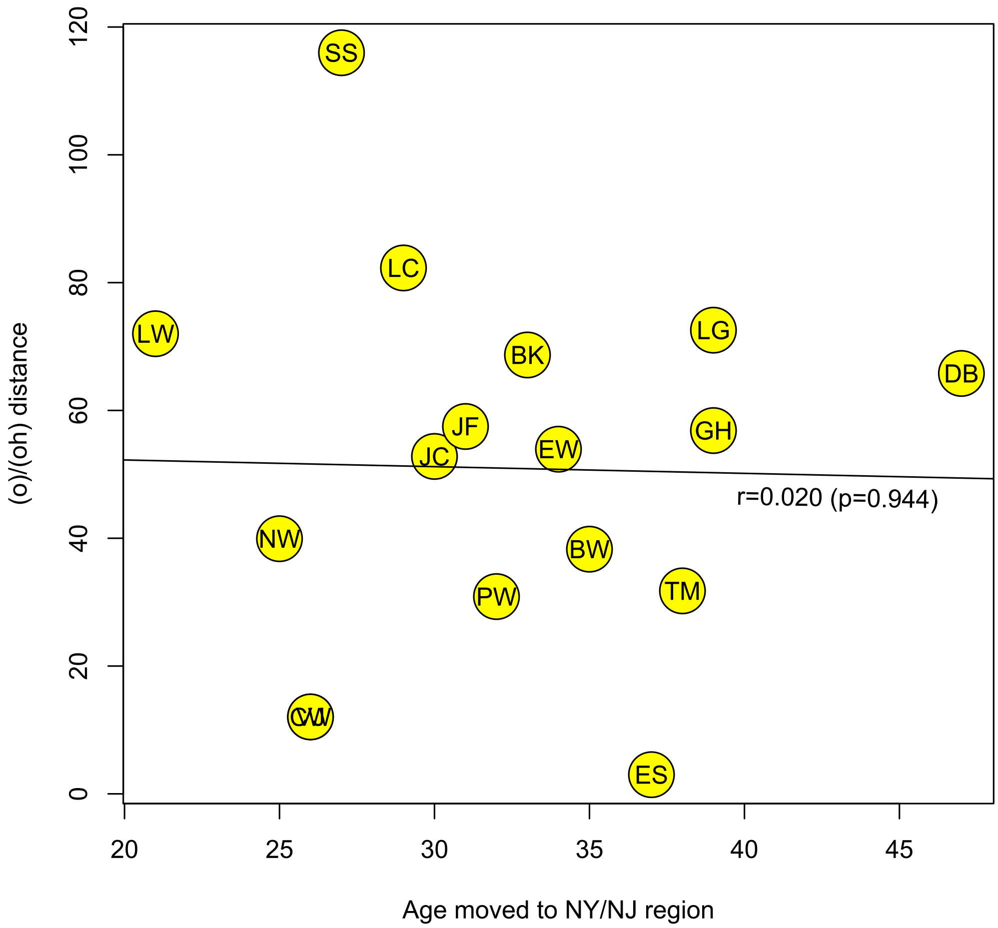


Figure 7.4: (o)/(oh) distance vs. age moved to NY/NJ

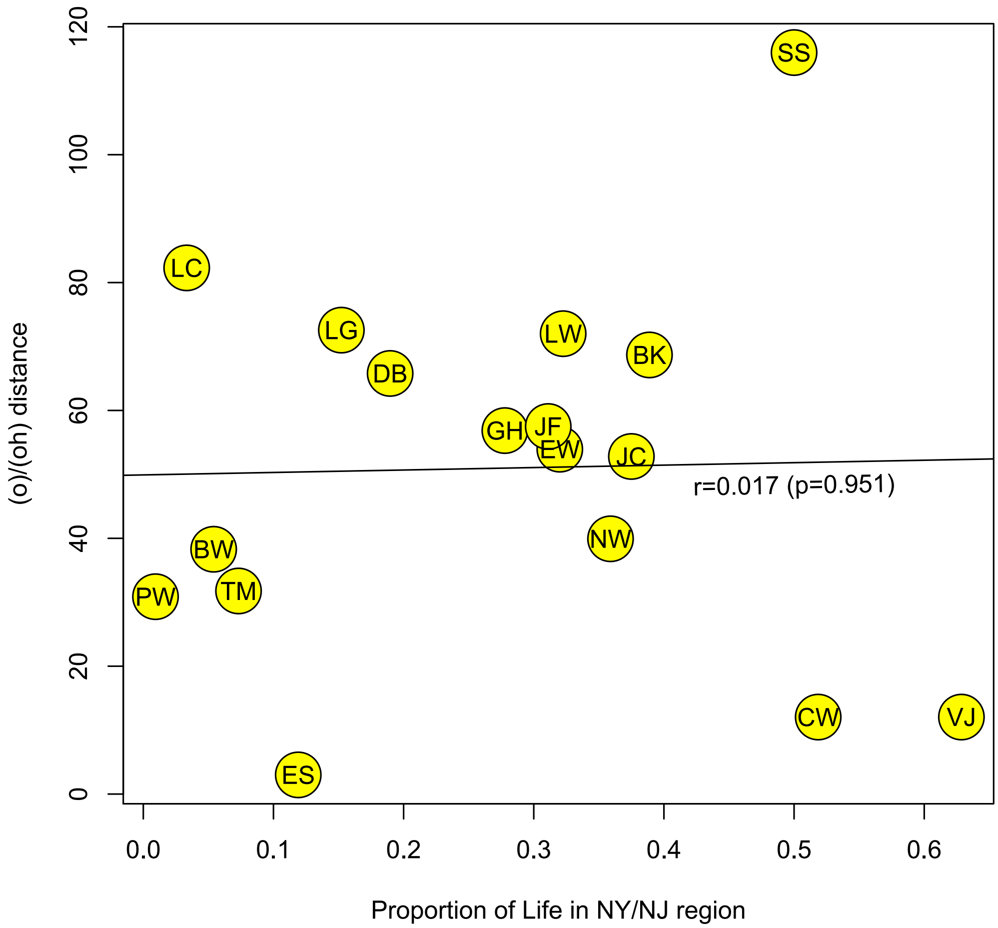


Figure 7.5: (o)/(oh) distance vs. proportion of life in NY/NJ

this measure still does not show any significant relationship with (o)/(oh) distance (r=0.017, p-value = 0.951)

In this group of speakers, then, there is no indication that time spent in the U.S. (whether absolute or proportional) has an effect on acquisition of the low back vowel contrast.

7.4.2 Time and (aw)-raising

An analysis of the effect of Time on Canadian Raising yielded more promising results. Fig. 7.6 shows raising in nonsalient (aw)T words plotted against absolute time spent in the New York region. There is no significant relationship between these variables (r=-0.25, p=0.337). Two well-behaved speakers at the margins strengthen the appearance of a negative relationship: LC, who had only been in New York City for a year at time of interview, shows a high degree of raising in this context, while 40-year Upper West Side veteran VJ does not raise. However, the remaining 15 speakers form a structureless blob. Similarly, there is no significant relationship between (aw)T-raising and Age Moved to the region (Fig. 7.7, r=0.37, p=0.142) or (aw)T-raising and proportion of life in the region (Fig. 7.8; r=-0.35, p=0.171).

Despite the lack of significance for any of these relationships, it is worth noting that the r values are much larger, and the p values much smaller, than those obtained in the comparisons between low back vowel distance and any of the temporal measures. Moreover, all three results trend in a sensible direction: greater absolute or relative time spent in the New York region is negatively correlated (albeit weakly, and not significantly) with degree of (aw) raising in (aw)T words, while Age Moved is positively correlated (again, weakly) with degree of raising. That is, there is some suggestion that the more time spent immersed in nonraising NYAE, the more attenuated raising becomes; conversely, the more time spent around CE, the more raising seems to be maintained.

The relationship between time and raising in *out* is even more suggestive. Fig. 7.9 show the amount of raising in *out* plotted against Years in the NY/NJ region; while the rho of -0.40 is not significant (p=0.116), there does appear to be a

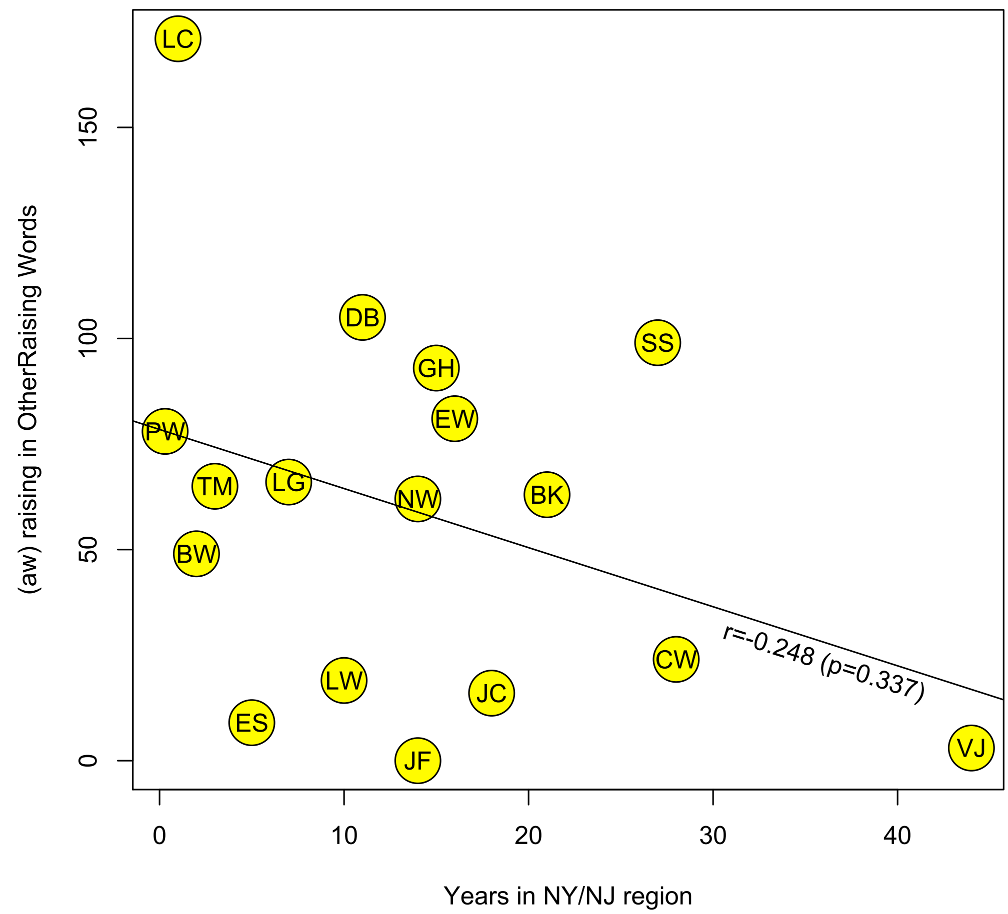


Figure 7.6: (aw) raising in AWT words vs. years in NY/NJ

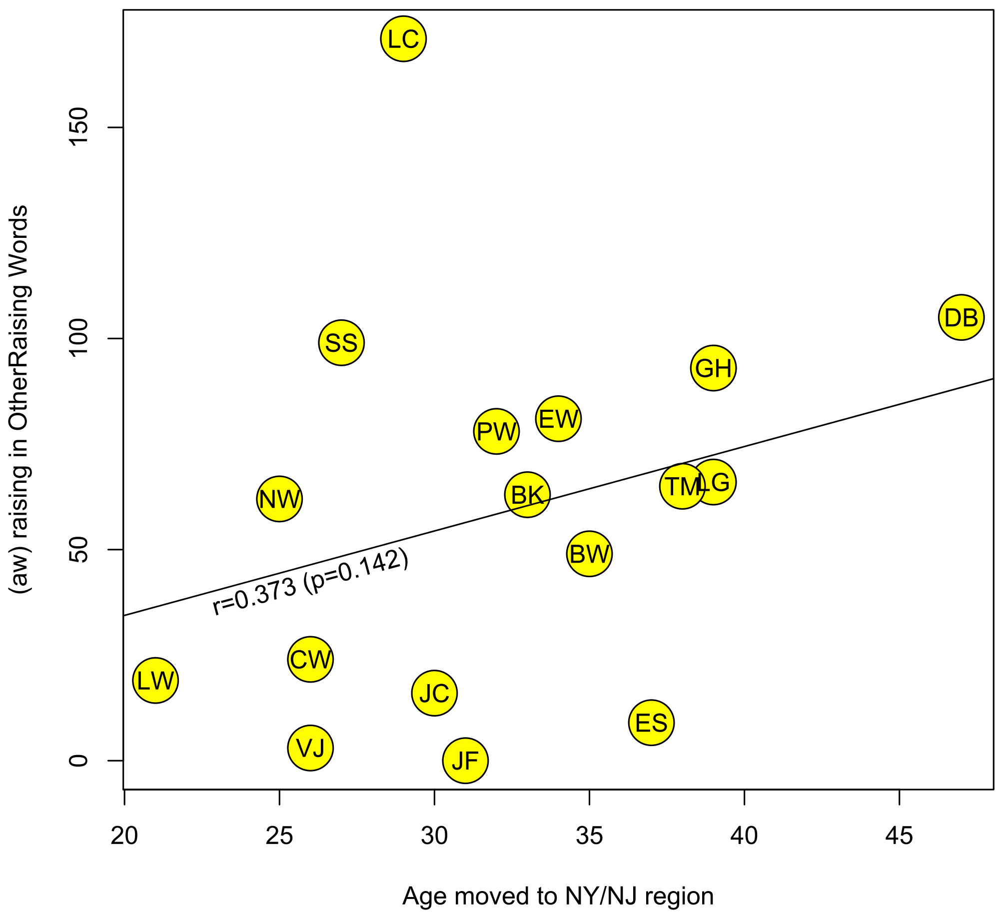


Figure 7.7: (aw) raising in AWT words vs. age moved to NY/NJ

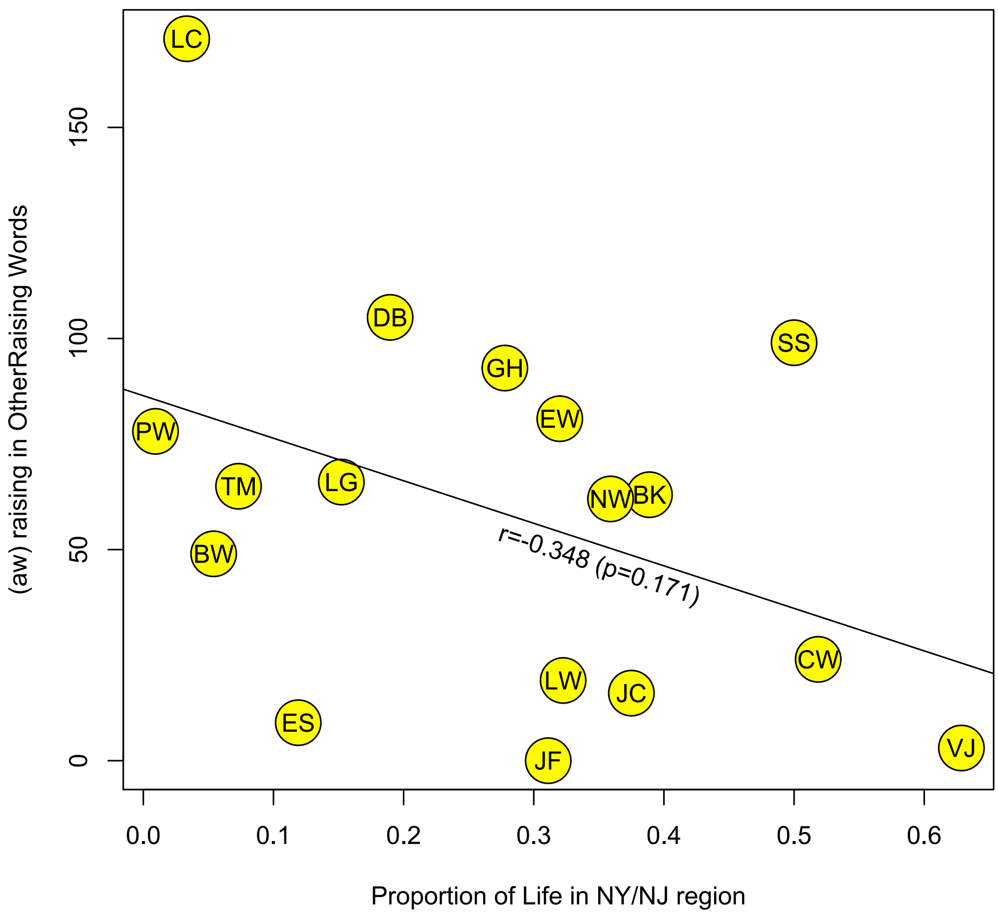


Figure 7.8: (aw) raising in AWT words vs. proportion of life in NY/NJ

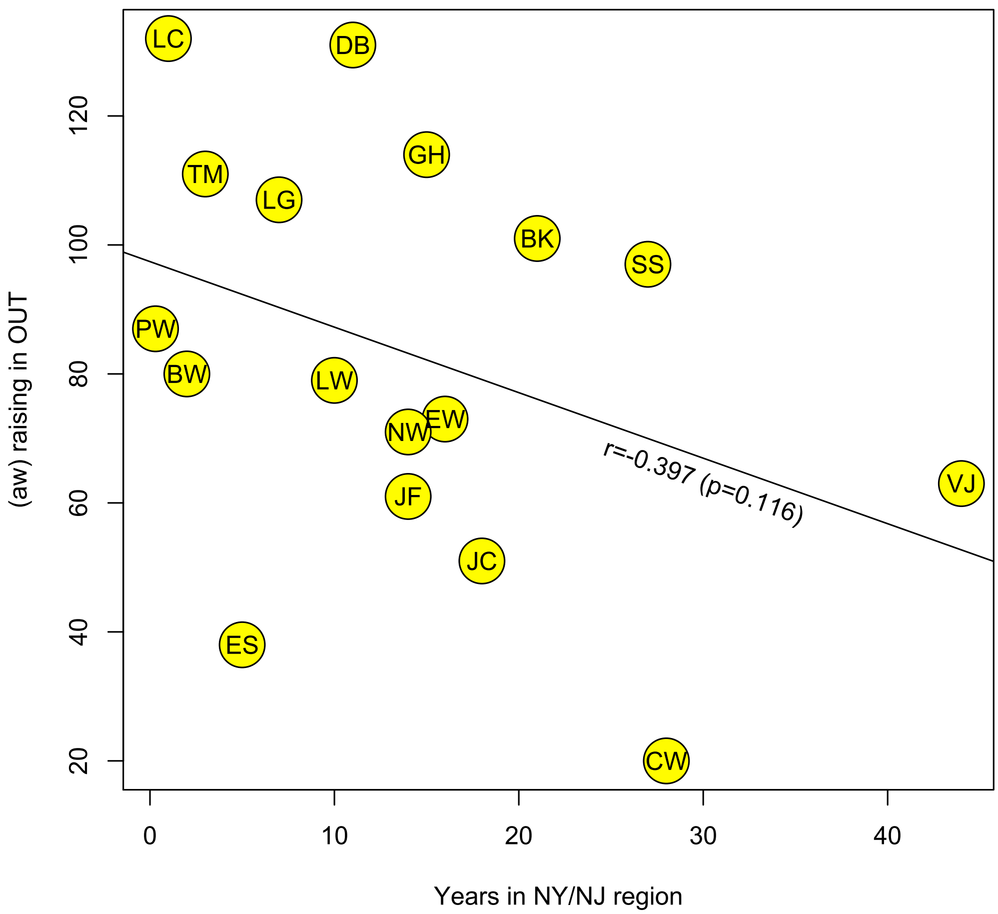


Figure 7.9: (aw) raising in OUT vs. years in NY/NJ

negative relationship between these features. The relationship between Age Moved and *out*-raising approaches significance (r=0.49, p= 0.050) (Fig. 7.10), as does that between *out*-raising and Proportion of Life (r=-0.47, p=0.060) (Fig. 7.11). Again, while the relationship between *out* raising and Time does not quite reach significance for any measure, the trends follow a sensible pattern: more time in the new region seems to mean less raising in *out*, while more time in the home region means more *out*-raising.

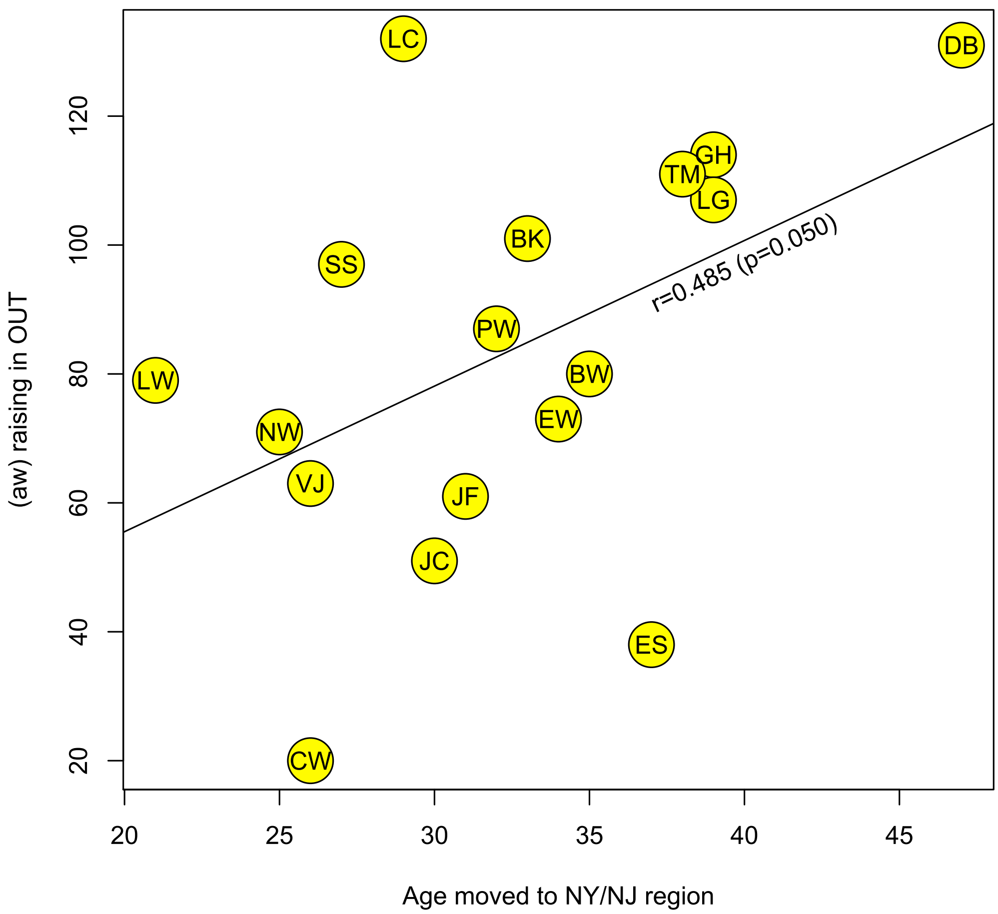


Figure 7.10: (aw) raising in OUTvs. age moved to NY/NJ

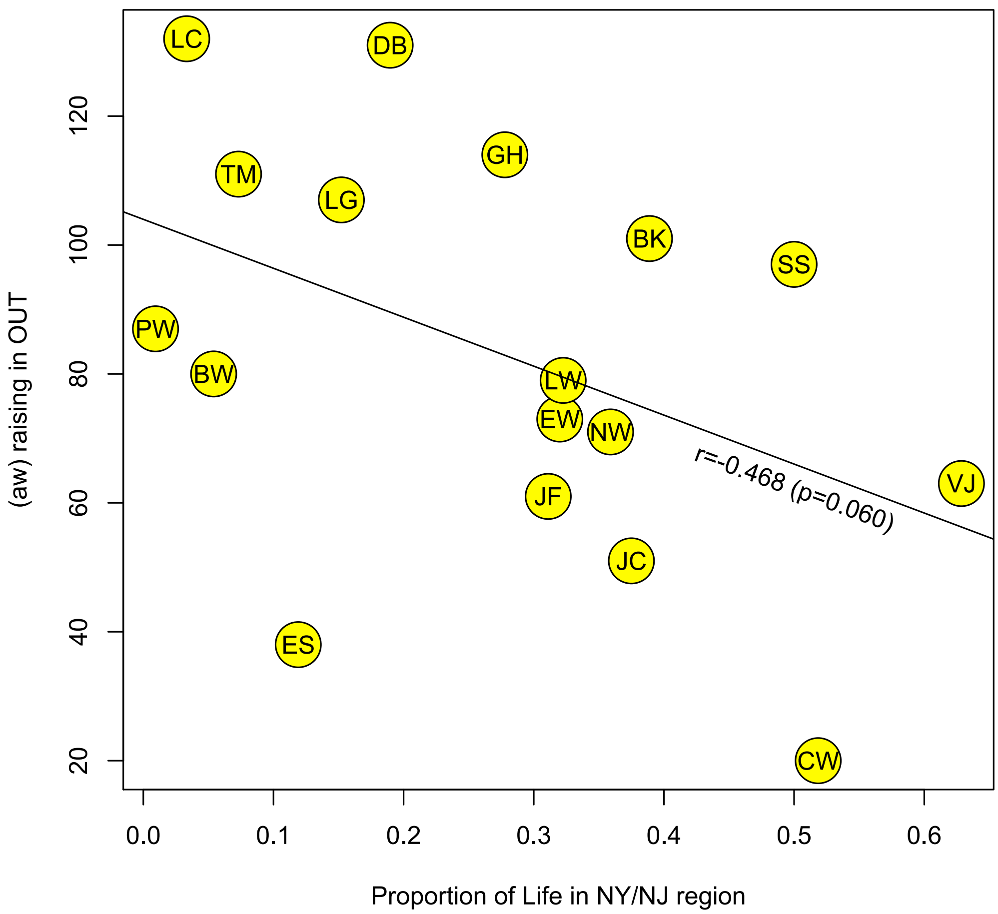


Figure 7.11: (aw) raising in OUT vs. proportion of life in the NY/NJ

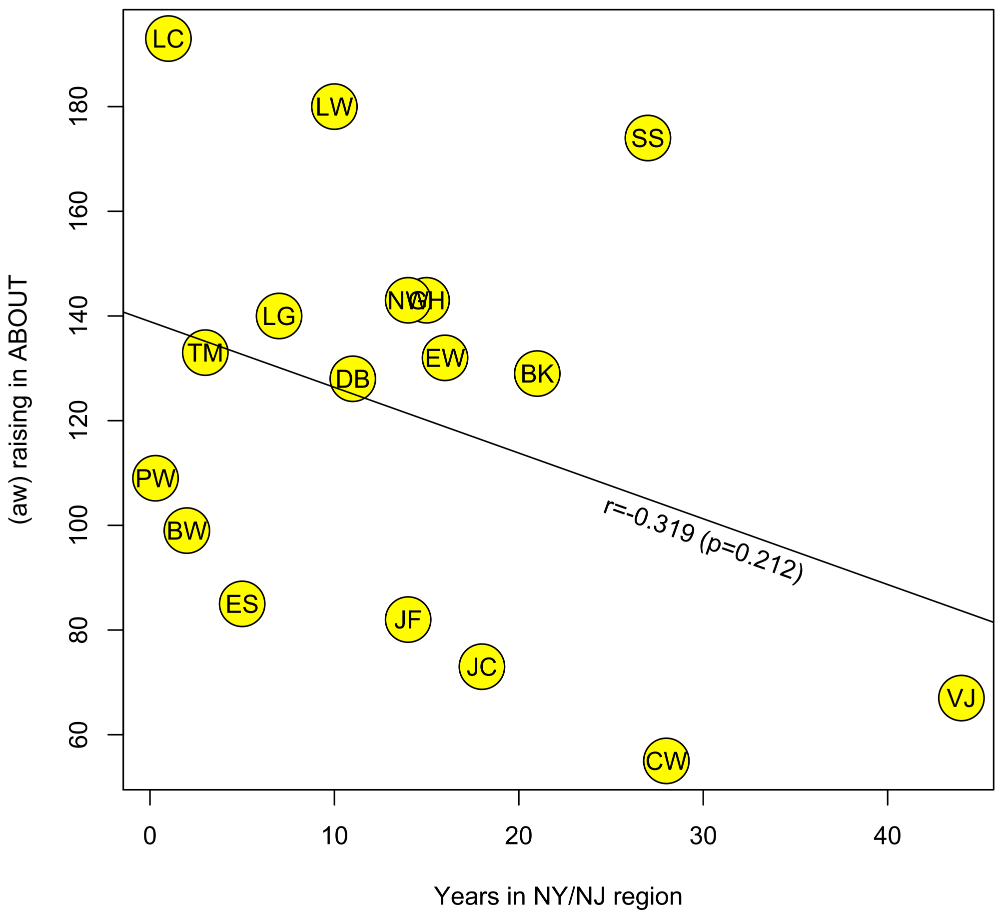


Figure 7.12: (aw) raising in ABOUT vs. years in NY/NJ

Finally, Figs. 7.12, 7.13 and 7.14, show the amount of raising in *about* plotted against Years in the NY/NJ region, Age Moved, and Proportion of Life in new region. Again, none of these factors are significantly correlated with *about*-raising, though there seems to be a suggestion of a negative relationship between time spent in the new region (relative or absolute) and degree of raising in this word.

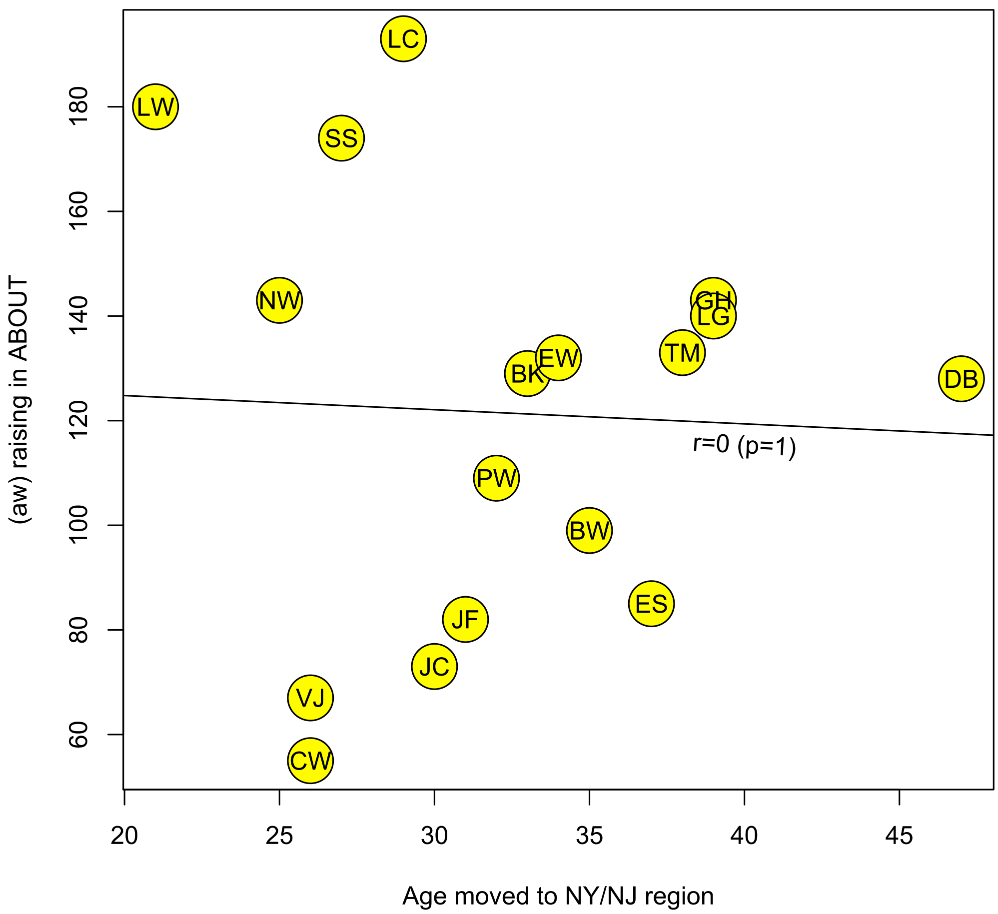


Figure 7.13: (aw) raising in ABOUT vs. age moved to NY/NJ

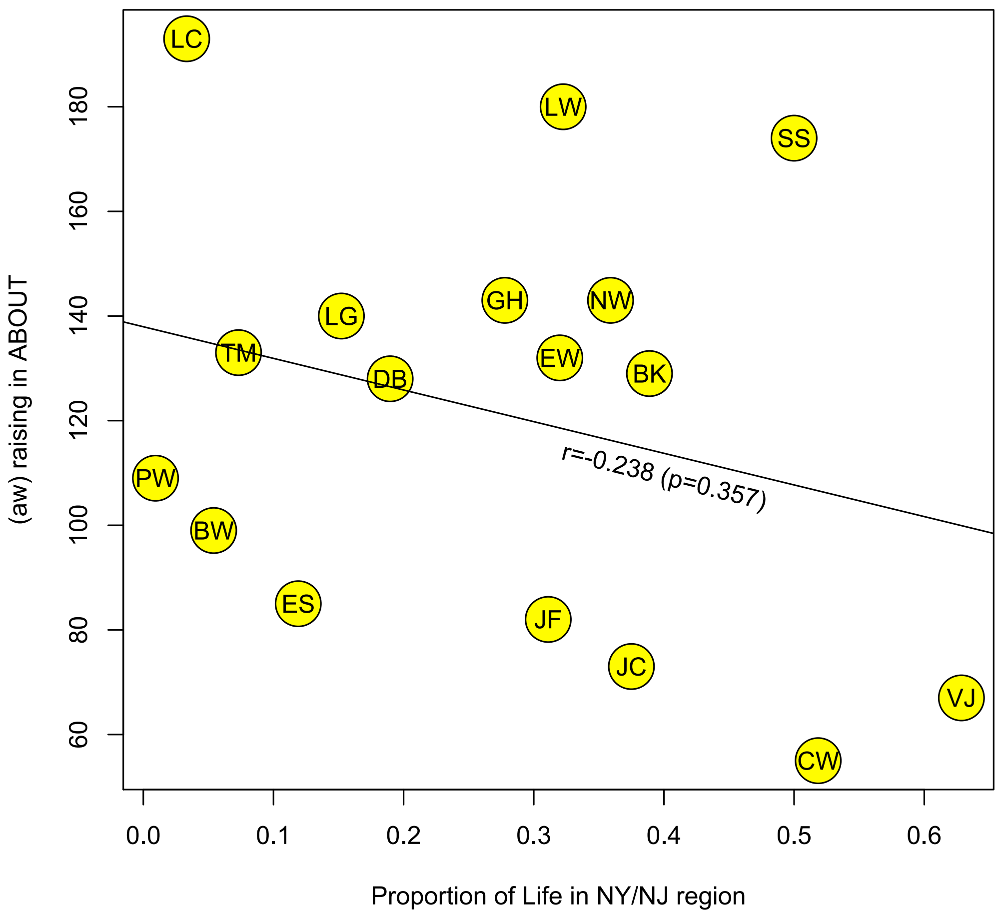


Figure 7.14: (aw) raising n ABOUT vs. proportion of life in NY/NJ

7.4.3 Summary of the effect of time on each feature

Neither low back distance nor degree of Canadian Raising shows a statistically significant correlation with any of the temporal measures of dialect exposure subject to analysis. However, there does seem to be a difference between these two features with respect to the effect of time. While low back vowel distance shows absolutely no evidence of covarying with time spent in either the new dialect region or the old region among these speakers, the results for (aw) raising do consistently trend in a particular direction, with increased time in the new dialect region associated with less raising across (aw)T words, *out*, and *about*, and increased time in Canada (as captured by Age of Move) associated with greater degrees of raising.

7.5 Interaction with speakers of the new dialect

Accommodation to a new variety requires not only time, but opportunity, in the form of exposure to interlocutors emitting tokens of new dialect variants. Research on language and social networks has shown that many and varied interactions with speakers of a variety will result in greater use of that variety (e.g. Milroy 1987). While detailed information about the social networks of the speakers in this study was not collected, one way to investigate the effect of exposure to speakers of a new dialect is to consider a single important interlocutor, such as the spouse or partner of a speaker. One's partner is likely to be a constant source of linguistic tokens, as well as someone to whom a speaker is very likely to accommodate (Zajonc et al. 1987).

Table 7.5 lists each speaker along with the home speaker of the speaker's partner (current, or if none current, most recent ex). The third column notes whether

Speaker	Partner from	(o)/(oh) different	(aw)-raising
BK	England (ex)	yes	no
BW	Canada	no	yes
CW	New Hampshire (ex)	no	no
DB	New York City	yes	no
ES	England (ex)	yes	no
EW	Taiwan	no	no
GH	Quebec	no	yes
JC	California	no	no
JF	Vietnam/California	no	no
LC	Canada	no	yes
LG	Binghamton, NY	no	no
LW	New Jersey	yes	no
NW	South Indian	no	no
PW	none	no	no
SS	New York City	yes	no
TM	South Carolina	yes	no
VJ	Canada (deceased)	no	yes

Table 7.5: Speaker partners: Their origins and inferred status with respect to (o)/(oh) and raising

each partner hails from a 2P, (o)/(oh) distinguishing dialect region, and the fourth column whether each partner has come from an (aw)-raising region (practically speaking, this column codes whether the partner is Canadian). The actual status of each partner's (o)/(oh) word classes and realization of (aw) are guesses in every case, as I did not interview or even speak with anyone's partner in the course of this study. However, it is an educated guess, insofar as we know the distribution of these features across dialects of English. If a partner is not a native speaker of English, it is assumed that he/she does not distinguish (o)/(oh) or raise (aw).

	Partner same	Partner different
Speaker same	CW, NW, PW, VJ	ES, TM
Speaker different	BW, EW, GH, JC, JF, LC	BK, DB, LG, LW, SS

Table 7.6: Partner inferred (o)/(oh) status vs. Speaker (o)/(oh) status

7.5.1 Partner and (o)/(oh) difference

Table 7.6 categorizes each speaker according to the dialect background of their partner (1P or 2P) and whether that speaker showed a significant difference between (o) and (oh) on either phonetic dimension. There seems to be no evidence in this data for a relationship between the low back vowel status of a speaker's partner and the speaker's own low back vowel status. The 11 speakers who make a contrast between these vowels are evenly divided between partner types. Among the 6 speakers who do not make a significant low back vowel distinction, 4 do not have exposure to a 2P partner, but this difference is hardly striking.

However, as seen in Chapter 4, speakers show much more fine-grained variation in terms of the actual distance between these vowels. Fig. 7.15 again reproduces the scatterplot of effect sizes, color coding each speaker according to the assumed low back vowel system of the partner. Unsurprisingly, there is no clear color grouping along the (aw) raising dimension. However, there seems to be a relationship with (o)/(oh) distance: 9 of the 11 speaker with less than 60 Hz difference between these vowels do not have a 2P partner, while 5 of the 6 speakers with more than 60 Hz difference do have a 2P partner. This difference is not significant,[4] but might be worth probing in future research.

[4]according to an ANOVA testing whether the two partner groups differ according to Euclidean distance.

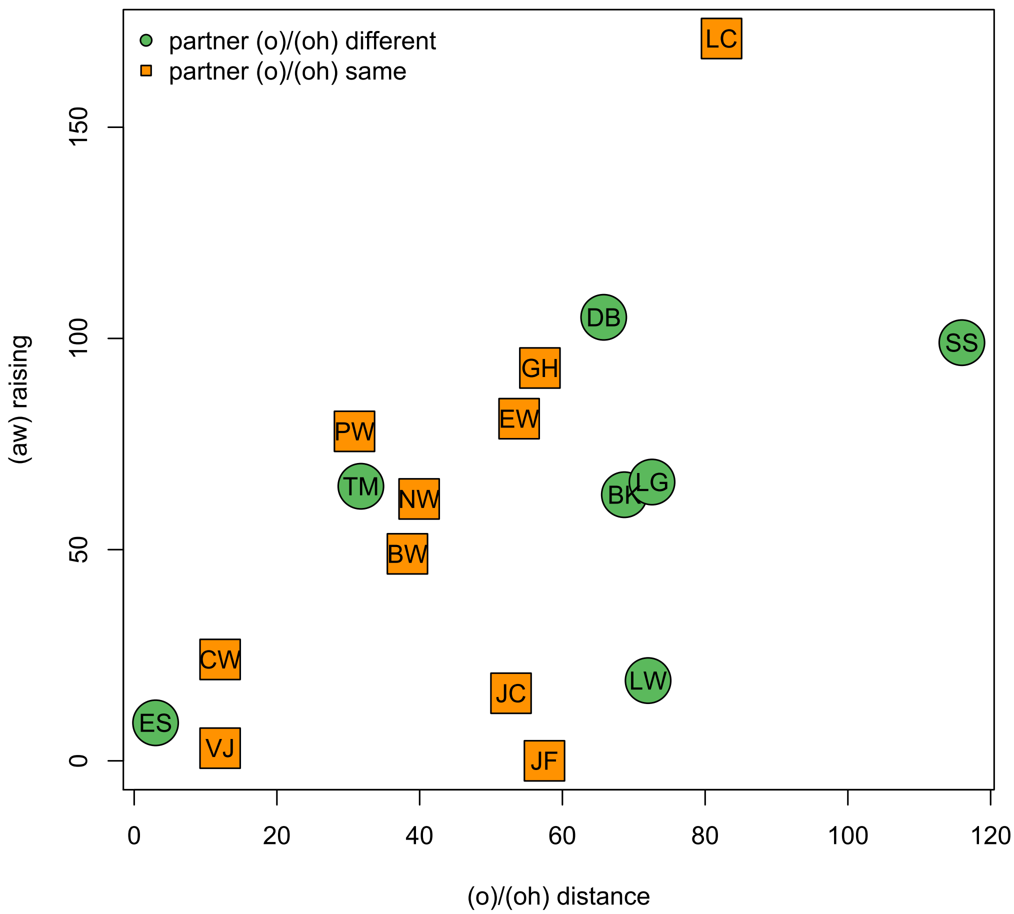


Figure 7.15: (o)/(oh) distance vs. (aw) raising, by inferred (o)/(oh) system of partner

	Canadian Partner	Non-Canadian Partner
Speaker raises	GH, LC	DB, EW, LG, PW, SS
Speaker does not raise	BW, VJ	BK, CW, ES, JC, JF, LW, NW, TM

Table 7.7: Partner Canadianness vs. Speaker (aw) raising

7.5.2 Partner and (aw)T-raising

Table 7.7 categorizes each speaker according to the dialect background of their partner (CE or non-CE) and whether that speaker showed significant raising in (aw)T words. There is no indication of a relationship between the qualitative (aw) raising behavior of a speaker and their partner's inferred raising status. The four speakers with Canadian spouses are evenly split between raisers and nonraisers. The speakers with non-Canadian partners are somewhat weighted towards non-raising, though this is not significant.

Unlike low back vowel distance, degree of Canadian Raising shows no indication of being affected by speaker partner. As Fig. 7.16 shows, the 4 speakers with Canadian partners are evenly spread out throughout the space for both variables.

7.5.3 Summary of partner effect on each feature

No significant effects of (inferred) partner low back vowel contrast or (aw)-raising status were found on the linguistic behavior of speakers in this study. However, the low back vowel system and Canadian Raising are again differentiated by their apparent relationship to Partner dialect: while there is no indication that a partner's (aw) raising status affects a speaker's degree of (aw)-raising, there is a strong

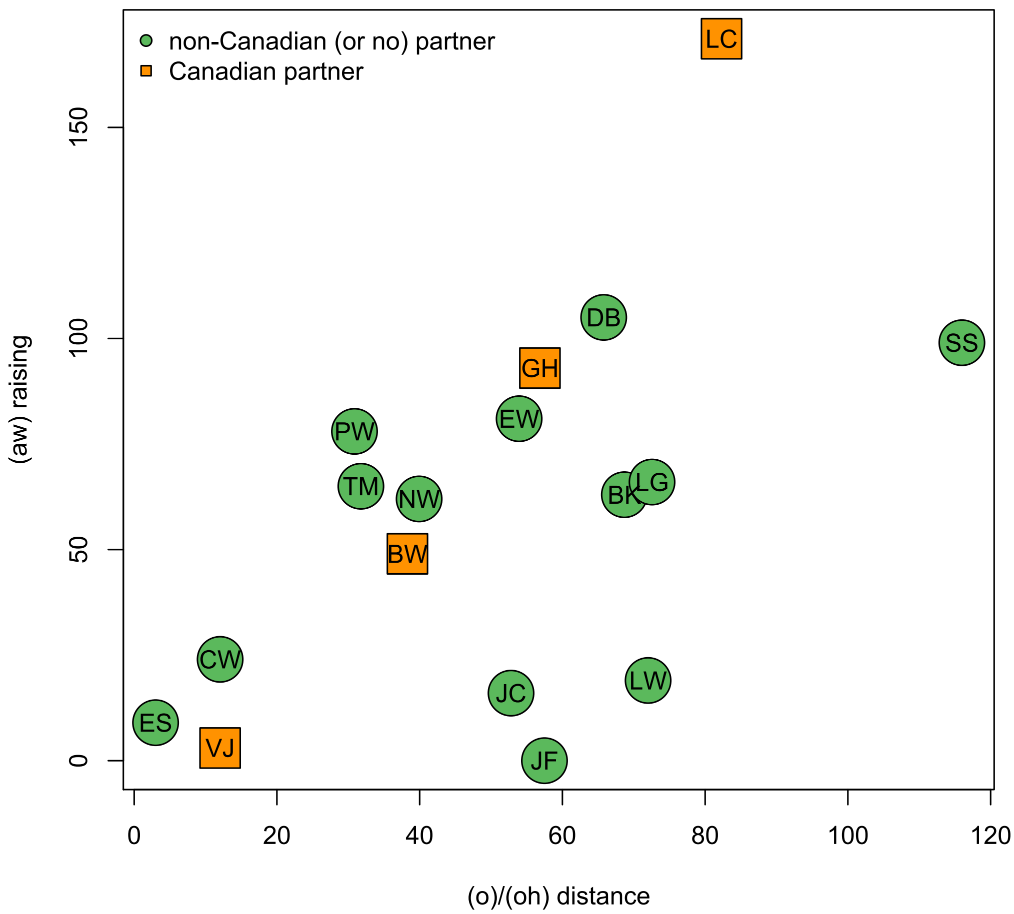


Figure 7.16: (o)/(oh) distance vs. (aw)T raising, by Canadianness of partner

suggestion that partner low back vowel status may affect low back vowel realization. Speakers with a probable (o)/(oh)-distinguishing partner seem more likely to have a 60Hz or greater difference between their own low back vowels, while those speakers without such a partner almost all have less than a 60Hz difference between these vowels.

7.6 Intent to remain in the New York region

Attitude towards a new region is another factor which may affect the extent to which speakers acquire features of a new dialect. A cooperative or otherwise positive attitude towards an individual interlocutor tends to foster linguistic accommodation; this effect seems to generalize to the group or region, such that speakers who feel integrated within and positive towards the larger speech community are more likely to adopt its linguistic patterns.

However, attitudes towards a place and its people can be complex. Because I did not formally measure attitudes in a standardized way in this study, it is difficult to objectively categorize these speakers in terms of their attitudes towards their home and adopted countries (or regions within). Each speaker might be coded as having a positive or negative attitude towards the U.S., for instance, based on my interpretation of statements made during the interview; unfortunately, hardly any speaker expressed a consistent, clearly negative or positive view towards either home or adopted region.

One thing that each speaker did express in clearer terms was his intention or desire to remain in the U.S. Intention to stay in the new dialect region is not a perfect proxy for attitude; it may reflect a speaker's economic constraints as much as his affection for a new city or country. However, it does indicate something

	(o)/(oh) same	(o)/(oh) different
Back to Canada	ES, PW	BW, DB, LC, LG, LW
Staying in NY/US	CW, NW, TM, VJ	BK, EW, GH, JC, JF, SS

Table 7.8: Intent to Stay in U.S. vs. (o)/(oh) status

	(aw)T raised	(aw)T lowered
Back to Canada	DB, LC, LG, PW	BW, ES, LW
Staying in NY/US	EW, GH, SS	BK, CW, JC, JF, NW, TM, VJ

Table 7.9: Intent to Stay in U.S. vs. (aw)T raising

about a speaker's overall attitude towards the new region ("Could I consider this place to be home for the rest of my life?") as well as motivation to accommodate.

Each speaker was therefore coded as *Staying in NY/US* or *Going back to Canada*, based on what they said about this subject during the interview. Table 7.8 shows how speakers are distributed by their desire to stay in the U.S. and their realization of the low back vowel contrast, and Table 7.9 shows desire to stay vs. presence of (aw) raising. No effect of intent to stay on either linguistic variable is apparent. This is also the case in the comparison of how Remainers and Returners pattern with respect to the quantitative realization of each variable (Fig. 7.17): the people who plan to return to Canada vary just as much with respect to both features as those who intend to stay.

Intention to remain or return, then, does not seem to have any effect on the realization of either variable for these speakers.

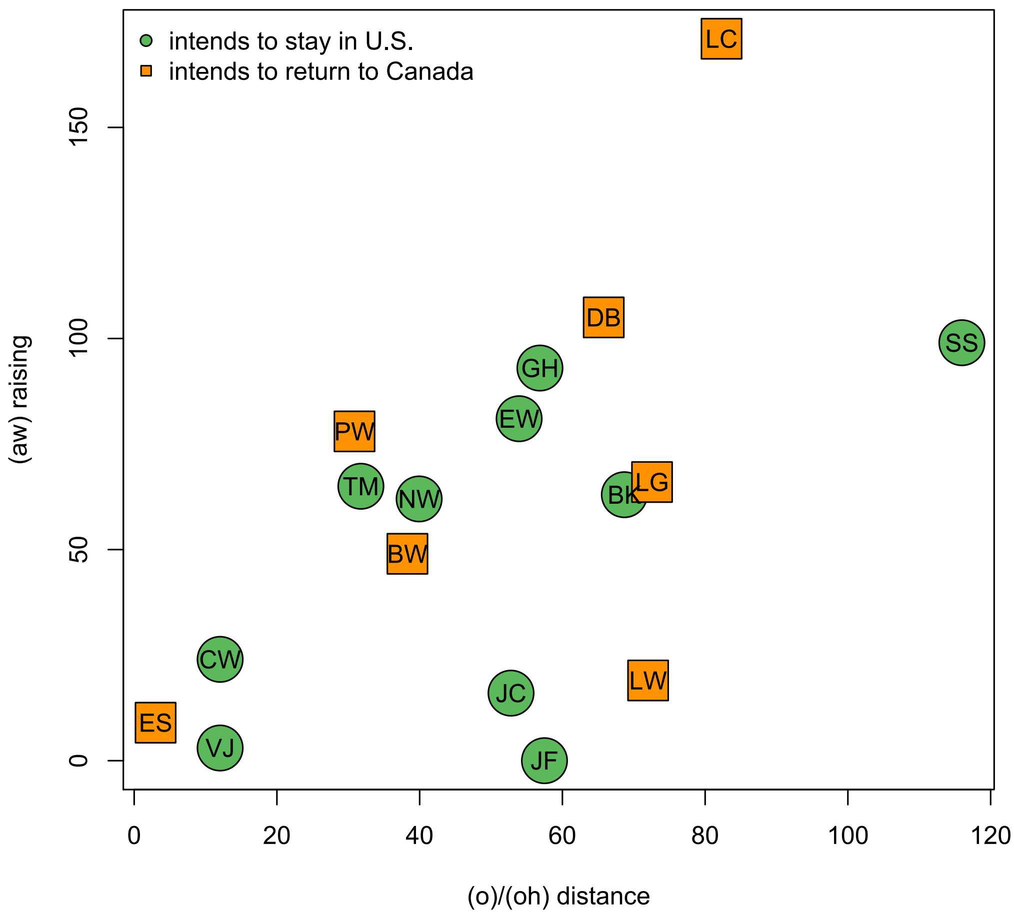


Figure 7.17: (o)/(oh) distance vs. (aw) raising, by intent to stay

7.7 Awareness of linguistic features

Another factor which may affect whether a speaker will accommodate to a given variable is the speaker's awareness that this variable differs across dialects. If a variable is above the level of conscious awareness, then the speaker may have more control over how this feature is used in speech. The result of this control, however, is hard to predict; awareness may hasten or retard accommodation, depending on whether the feature is positively or negatively viewed.

The two features in this study differ markedly in the extent to which speakers are aware of them. Every speaker in this group expressed awareness of Canadian Raising at some point during the course of the interview, either on their own, or when asked about the features of language which differ between CE and the local variety. This reflects the broader salience of this feature as a marker of CE among North Americans: Americans and Canadians alike know that Canadians raise their diphthongs in *out*, *about* and other (aw)T words.

The (o)/(oh) merger (or lack thereof), however, is a comparatively stealthy dialect feature; this merger, like most mergers, is below the level of conscious awareness. However, by explicitly probing awareness of this distinction using the minimal pair and rhyming pair tasks, I found that some speakers in the group were aware of this distinction in NYAE. Does such conscious awareness of the distinction at all affect whether speakers realize the distinction in their own speech?

7.7.1 Minimal Pair and Rhyming Pair awareness results

As described in Chapter 2, speakers completed an Other Dialect Judgment task in which they were asked to consider how speakers of the new local dialect produced various words. After completing the Minimal Pair and Rhyming pair tasks in the standard way, speakers were asked to look back over the list of pairs they had just read and evaluated, and say whether they thought people from the New York region would either say any of the words on these lists differently or have different judgments regarding whether certain pairs sound the same or different. Of course, speakers commented on various items on this list that had nothing to do with either of the features of this study; for instance, many speakers noted that word pairs with coda (r), such as *barn/born* and *fire/higher*, might be produced without the (r).

However, the main point of this exercise was to probe the low back vowel distinction. As described in Chapter 4, in their first, typical pass though the Minimal and Rhyming Pair tasks, the speakers in this study were nearly all merged in both production and perception, reflecting the norm of their native dialect, CE. The purpose of the Judgment Task was to determine whether any of these speakers had explicit awareness of the new dialect norms: that is, do they know that speakers of NYaE make a distinction between (o) and (oh)?

The results of this task were not always conclusive. However, there are some speakers who clearly grasp that there is a (o)/(oh) distinction in the second dialect and some who are completely unaware of this difference.

Seven speakers display a strong awareness that there is a contrast between these two vowels, as well as an accurate, if exaggerated, grasp of the nature of the phonetic difference. GH, JC, JF, LW, NW, and TM noted the difference for many of the (o)/(oh) pairs on the list, producing an extremely high back (and often lengthened) vowel for the (oh) word in each pair. LC is also aware of the contrast, but varies in where she locates the difference between CE and NYaE. For instance, she notes that *caught/cot* are different, saying that *caught* sounds like [kʊᵊt], but for *don/dawn* and *odd/awed* said that the difference is due to*don* and *odd* being produced, respectively, [dan] and [ad], with a very fronted low vowel. These 7 speakers also made statements indicating awareness of a more general contrast beyond the individual differences between words on this list; these generalizations usually referenced orthography, e.g. LC's observation that "a lot of the times just in general *o*'s are *a*'s, like *dot com* is [dat kam], like it's an *a* sound."

LG, interestingly, is aware of the difference, but for the most part gets the phonetics wrong, saying that *talk* and *caught* are produced by locals as [tak] and [kat]. However, she does note that New Yorkers say *dog* like [dʊᵊg].

CW's responses are more difficult to interpret. The only pair she says would be different for New Yorkers is *caught/cot*, and produces the right phonetic distinction, with *caught* having the higher backer realization. However, for the remainder of the pairs, she attempts both words with the exagerrated high back vowel, then the lower fronter vowel, before deciding that they are probably the same. A possible interpretation of this behavior is that while she is unaware that there is a general contrast, she does grasp that there is a wider range of acceptable pronunciations for this putatively single vowel category.

Four of the speakers pick out 1 or 2 words or word pairs as being different, but do not show conscious awareness of a general contrast. BW, given the (distractor) pair *coal/call*, says people from Jersey say [kwal], and points out a "subtle elongation" of the vowel in *pawned* as compared with that in *pond*, but otherwise does not seem to generally grasp that there is a difference. PW says *caller* is more "drawn out" than *collar*, but produces the first word with a much more fronted vowel. SS says *tall* may be different from *doll*, but doesn't point out any other pairs. VJ says *doll* may be produced with a fronter, more "drawn out" vowel, but otherwise does not spot any low back differences.

Finally, BW, DB, ES, and EW betray no awareness of a difference in the low back vowels, either phonological or phonetic. In summary, 4 speakers seem to be clearly unaware of the low back vowel contrast, 7 speakers appear to have an accurate grasp of the general contrast as well as its phonetic realization, and the remaining 6 speakers fall somewhere in between.

7.7.2 Awareness of contrast versus realization of contrast

Having established that there is variation in awareness of the low back vowel contrast in NYaE, I then investigated whether there is any relationship between awareness and realization of each of the features in my study.

Unsurprisingly, there is no apparent effect of awareness of the low back vowel contrast on realization of Canadian Raising; perhaps more surprisingly, there is also no affect of awareness of this contrast on realization *of the contrast*. Fig. 7.18 color codes speakers by awareness level. The speakers who show the clearest awareness of a general contrast in NYaE cluster in the middle of the distance range for (o)/(oh), indicating that awareness of this feature has neither an inhibitory effect on contrast nor an encouraging one.

This lack of relationship is also evident at the qualitative level. Table 7.10 divides the speakers into 6 groups, based on demonstrated awareness of the low back vowel contrast and production of this contrast. There is again no strong indication that awareness is linked to realization of the contrast; it so happens that 4 out of 6 aware speakers realize the contrast, but so do 3 of the 4 unaware speakers.

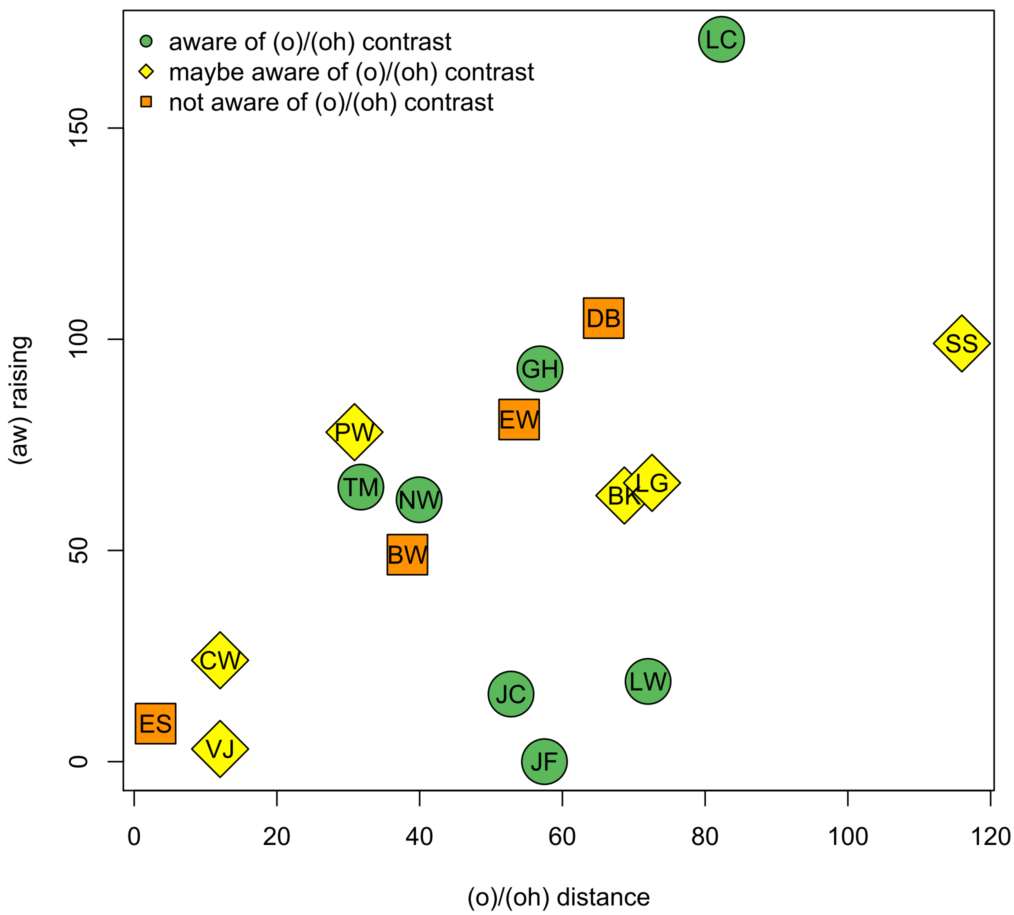


Figure 7.18: (o)/(oh) distance vs. (aw) raising, by awareness of (o)/(oh) contrast

Awareness of contrast	(o)/(oh) same	(o)/(oh) different
Unaware	ES	BW, DB, EW
Maybe aware	CW, PW, VJ	BK, LG, SS
Aware	NW, TM	GH, JC, JF, LC, LW

Table 7.10: Awareness of (o)/(oh) contrast vs. (o)/(oh) status

7.8 Conclusions

This chapter examined several extralinguistic factors for their possible effects on accommodation towards either the low back vowel distinction or the nonraised (aw) of NYaE.

Two of these factors, **Region of Origin** and **Gender**, reflect basic demographic information about the speakers. While there was no particular reason to expect that speakers would pattern differently depending on these factors, these factors are often relevant in language studies. There was some suggestive patterning found for both factors: specifically, speakers hailing from Western Canadian provinces and males seem to cluster on the low-distance end of the low back vowel distance range. However, members of both of these groups among these speakers also tend to have partners who do not contrast (o)/(oh).

One factor, **Intent to remain** in NY/the U.S, taps into speaker attitudes about the new region. One might expect that an expressed intent to remain in the new region would predict accommodation to NYaE realizations of both low back vowels and (aw); however, there seems to be no relation between this factor and these features in this group of speakers.

Finally, I examined whether **Awareness** of low back vowel contrast in the new region affects speaker realization of this contrast. Again, there is no indication of a relationship between this factor and realization of either feature.

The two extralinguistic factors which do seem to be related to use of these features are also the two factors which relate to new dialect exposure: **Time** and **Partner dialect**. These factors measure exposure in different ways: the temporal measures of exposure count years, while **Partner dialect** reflects something like intensity of exposure. Interestingly, these different types of exposure have different relationships to the features studied here. (aw) raising seems to decrease with time spent in the new region, but is unaffected by partner dialect. Low back vowel contrast, meanwhile, may be affected by partner dialect, but bears no relation to time spent in the new region.

A possible reason for this difference might be the differing status of the two features among *speakers who happen to be in the NY/NJ region*, as distinguished from *speakers of* NYaE. The tri-state region is home to native New Yorkers and New Jerseyites as well as millions of transplants from all over the country and world. Many of these non-local speakers will realize a (o)/(oh) distinction in their own speech, but many will not distinguish these word classes.[5] The vast majority of non-local speakers will be similar to native NYaE speakers, however, in producing a low (aw)T. Therefore, while increased time in this region will result in a consistent gradual accrual of (aw) tokens which are overwhelmingly low, the input regarding the low back vowels may be more ambiguous; in such an environment, it may require the focused, consistent input of a single salient speaker to more fully realize the low back vowel distinction.

This difference aside, however, the fact that exposure seems to be the most important factor contributing to accommodation is consistent with the usage-based account of how individual change towards the new local features should occur: more exposure means more storing of local tokens, which leads to more shift in production.

[5]Moreover, those that do realize this contrast are not likely to realize it in the same way that speakers of NYaE do

CHAPTER 8

CONCLUSIONS AND FUTURE RESEARCH

8.1 The goals of this dissertation

As long as sociolinguistics has existed, scholars working within this field have argued that variationist data is relevant to, and indeed crucial for, building a linguistic model that adequately accounts for speakers' knowledge of and use of language. Often, this argument has focused on the integration of the social with the linguistic, drawing on evidence that speakers have knowledge of how different variants are socially indexed and use this knowledge in both production and perception; Eckert (2000), for example, discusses how adolescent speakers in a Detroit suburb use increased centralization of the nucleus of /ay/ to index urban orientation, non-conformist stances, and topics associated with adolescent rebellion, while Niedzielski (1999) has shown that speakers' assumed nationality (U.S. or Canadian) affects how vowels are perceived. However, because such studies are usually designed to uncover social patterns and effects, they are limited in terms of their specific impact on questions of phonological theory (beyond, of course, providing evidence that social factors must be incorporated into this theory). Other work focuses instead on accounting for quantitative variation within particular theoretical frameworks, usually by generating precise predictions from different accounts and then seeing whether these predictions are borne out in existing corpora; Guy (2007), for example, does this in his discussion of the "phonological" versus "lexical" approaches to lexical exceptionality. Such studies do an excellent job of addressing theoretical questions, but often leave social factors by the wayside.

This study has fruitfully combined these perspectives, by taking naturalistic sociolinguistic data and analyzing it in such a way that specific linguistic predictions can be tested. Chapter 2 described the group of speakers that were subject to study: native speakers of CE who moved to the New York region as adults, and who are therefore in a position to acquire features of a second dialect of English. These speakers participated in a sociolinguistic interview and reading tasks that probed their personal experiences of living in Canada and the U.S., their attitudes towards these regions, their awareness of the linguistic features which differentiate their dialects, and whether they thought they had "become locals", either culturally or linguistically. In addition to the qualitative social data gleaned from the interview content, large amounts of quantitative linguistic data was also collected.

Chapter 3 laid out the phonological predictions regarding second dialect acquisition from two theoretical perspectives: Generative Phonology and Usage-based Phonology. The chapter began with a discussion of how each of these theories treats some of the major concepts in phonology, notably *contrast* and *generalizations.* While generative phonology is characterized by abstract, singular lexical representations that are separate from rules, usage-based phonology usually includes myriad phonetically-detailed representations which give rise to both individual word productions and higher-level generalizations. These differences between the theories lead to different predictions regarding how contrast and generalizations ought to change over time in the minds and speech of the individual; the remainder of the chapter describes these predictions for the two variables which are the focus of this study, (o)/(oh) contrast and Canadian Raising.

8.2 The findings

This study presented several linguistic findings about speaker's individual and group productions of low back vowel contrast, (aw)-raising, and both of these features considered together.

One general result which emerges is that these speakers show remarkable first dialect stability, even after many years of new dialect exposure. In minimal pair contexts probing the (o)/(oh) distinction, speakers continue to merge these vowels in both perception and production. In conversational speech, nearly all speakers continue to raise (aw), at least in salient words like *out* and *about*. Both behaviors reflect the norms of CE which the speakers in this study acquired as their first dialect.

However, there is also ample evidence that these speakers have been phonetically affected in their production of both of these features by their residence in the New York region and exposure to local speakers.. Chapter 4 showed how most speakers exhibit a small but significant phonetic difference between (o) and (oh) in conversational speech after phonological factors have been accounted for; across speakers, frequency effects consistent with a lexically gradual divergence of these two vowel categories were found, as predicted by usage-based theory.

Chapter 5 described how, while nearly all speakers continue to exhibit raising in *out* and *about*, several have curtailed raising in less salient raising context words; across speakers, there are again frequency effects which are consistent with a lexically gradual lowering of this group of words.

In summary, a picture emerges of change in both features, in a manner consistent with a usage-based account of change over the lifespan: both features of NYaE are being accommodated to in a phonetically and lexically gradual manner.

However, Chapter 6, which compared behavior with respect to both of these variables across speakers, turned up an unexpected finding: distance between low back vowel word classes and degree of (aw)-raising are positively correlated. This is surprising, as the heretofore supported usage-based account of accommodation seems to predict that these measures would be inversely related: as a speaker gains more exposure to tokens of the new dialect, low back vowel distance should increase, while raising should decrease, reflecting accommodation towards both features of the new dialect. Instead, those speakers who show the greatest amount of accommodation towards the second dialect feature of low back vowel contrast are also those which most strongly maintain the first dialect feature of Canadian Raising.

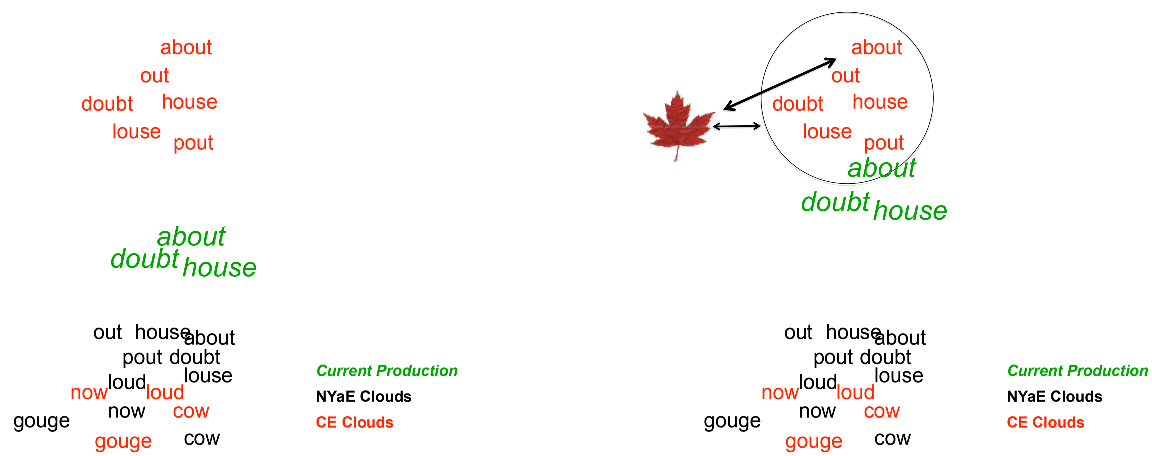
8.3 Explaining differential accommodation

So how to explain the puzzling behavior uncovered in Chapter 6? As described above, at first glance there does not seem to be a ready explanation in terms of the usage-based theory which had otherwise been supported by the independent results for low back vowel contrast and (aw)-raising. However, this is because the theoretical discussion thus far has included only two elements: stored exemplars and the linguistic category labels (here, words) attached to them.

According to formulations of usage-based theory in linguistics (e.g. Johnson 1997, 2006, Pierrehumbert 2006), exemplars may also be associated with a range of social labels, at varying levels of specificity: for example, an exemplar associated with the word label *dog* can also be tagged with the labels *Jen*, *female*, *American*, *young*, or with labels corresponding to any other salient social characteristic. When a particular social label is activated, via either external contextual factors or through internal factors (for instance, speaker desire to convey a particular social identity), exemplars associated with that label are more likely to be selected as models for production. This, of course, has been one of the attractive features of usage-based accounts for variationists and other linguists who seek a model which relates the social and the linguistic.

This social labeling is also crucial to understanding the relative patterning of low back vowel contrast and Canadian Raising in this data. As noted in previous discussions of these variables, they may be thought of as differing in terms of their linguistic status: in generative theory, (o)/(oh) contrast is a matter of phonological representation, while Canadian Raising involves a rule. Extralinguistically, these features also differ in terms of their social indexing: Canadian Raising is very strongly associated with (speakers of) CE, while low back vowel status has nowhere near this level of social salience. This is true for the speakers in this study, who unanimously point out that "*oot and aboot*" is a feature of CE, but only sometimes, with minimal pair-based prodding, note that NYAE speakers produce certain pairs of words differently; this asymmetry also characterizes North American speakers more generally. This difference in social indexing would be formalized (to the extent that anything is formalized) in a usage-based theory via a difference in social labeling of the relevant exemplars: raised tokens of (aw) are tagged *Canadian*, but tokens of (o)/(oh) words would not receive an equivalent social label.

This differential social labeling brings with it implications for how speakers should behave in various contexts, and in particular this study. Each speaker participated in a conversation about growing up in Canada, moving to the U.S. and living there as an expatriate, aspects of Canadian speech versus local American speech, and other topics surrounding being Canadian in a new country; the interview context was, essentially, a 90-minute activation of Canadian identity, and presumably, Canadian-labelled exemplars. Viewed in this way, it is not surprising that high levels of (aw)T raising were observed in this study: speakers were, to a certain extent, responding to the identity-heavy context in an appropriate way, favoring Canadian-labelled (aw)Ts which would tend to be located in the more raised region of the parametric phonetic space (Fig. 8.1) . Productions of (o) and (oh), meanwhile, would not be expected to show this identity-label-induced skewing: because no particular region of the phonetic space is likely to be favored for social reasons, tokens produced during the interview will in aggregate reflect the overall divergence of these categories which has occurred over time (Fig.8.2).



(a) Unmarked conversation

(b) Canadian-evoking conversation

Figure 8.1: Representation and Production of (aw) in the parametric phonetic space of a Canadian in New York. In unmarked conversation contexts, exemplars of (aw)T acquired from both CE and NYaE contribute to the production of new tokens of (aw)T, resulting realizations that fall somewhere between those of these two dialects (a). In contexts where Canadian identity is activated, exemplars associated with this label exert a stronger effect on production (b).

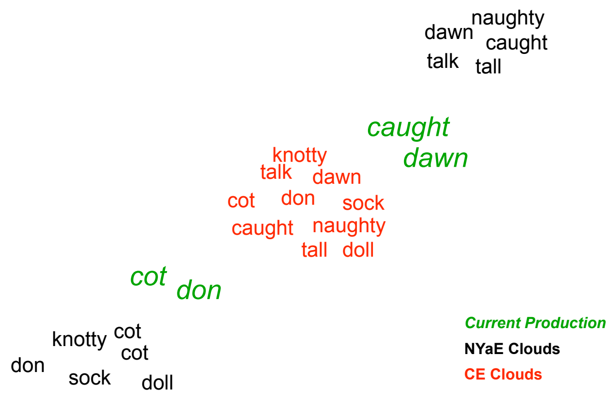


Figure 8.2: Representation and Production of (o)/(oh) in the parametric phonetic space of a Canadian in New York.

The inclusion of social labels enables an account of the overall group result seen in the expatriate Canadian data: evidence of low back distinction, but also evidence of Canadian Raising. This account also generates potentially testable predictions. For instance, it predicts that Canadian Raising should show much more variation in the speech of these individuals, depending on context: if these speakers were to be recorded in a more identity-neutral context that does not evoke Canadianness, everyone's levels of (aw)-raising should decline, while the extent of low back vowel distinction should remain stable.

However, this account is not yet complete. While it explains how (o)/(oh) distinction and strong Canadian raising may co-exist within a speaker, the positive correlation between these features across speakers requires a further addition to the model.

8.4 Proposal: Linguistic sensitivity

What factor relates a speaker's ability to acquire a (o)/(oh) difference with her tendency to use identity-appropriate tokens in a given speech context? It is difficult to find the answer to this question within linguistic theory. The broader field of theoretical linguistics is concerned with language universals, and sociolinguistics tends to be concerned with broader social patterns; even the '3rd wave' of sociolinguistics, with its focus on ethnography and individual experience, ultimately seeks to explain differences between individuals using more general notions of identity, within the social context. For these reasons, existing models of language use incorporate linguistic and social factors, but residual idiosyncratic variation displayed between individual speakers which cannot be explained via linguistic principles or social factors remains unaccounted for.

However, research in other areas suggests that this idiosyncratic residue is not mere noise, but at least partially explainable in terms of measurable differences in psychological characteristics. Differences in cognitive styles along many dimensions affect behavior in various domains: for example, differences in information-processing styles has been shown to affect judgmental behavior (Malhotra et al. 1983), differences in attachment styles affect how students approach studying (Gore and Rogers 2010), and differences in extraversion and other personality traits are related to the likelikhood of burnout among counselors caring for terminally ill patients (Bakker et al. 2006).

Similar effects have been shown to hold for linguistic behavior as well. For example, differences in learning style among young children can result in qualitatively different trajectories of language development, as described by Bates et al. (1988). More recent work by Yu (2010) has shown that individual differences in cognitive processing styles can affect the extent to which speakers perceptually compensate for coarticulatory effects, a finding with obvious implications for the way in which speakers perceive units of speech, store exemplars, and later produce the relevant categories.

To complete the account of D2 acquisition described above, I propose an additional component to the model, which I will refer to as *linguistic sensitivity* (*LS*). This idea is strongly influenced by the notion of "conversational sensitivity" described in e.g. Daly et al. (1987), Daly and Diesel (1992). According to Daly and Diesel, p. 412, "the underlying notion is that people differ in their sensitivity to what happens during conversations. It measures the propensity of people to attend to and interpret what occurs during conversations". Here, Daly and Diesel (1992) are referring to conversation in the macro sense: aspects such as tact, detecting multiple meanings in what people say, and perceiving nuances of affinities and other relationships conveyed through conversation.

LS, proposed here, incorporates these macro-aspects of conversational sensitivity while also including sensitivity to the patterning of smaller units which linguists are accustomed to studying; speakers may vary in their sensitivity to nuances of relationship as well as use of socially indexed linguistic variables or fine-grained phonetic variation. *LS* would not constitute an object of the usage-based model in the way that an exemplar or a category label does; rather it would be incorporated as a parameter of the model. Set the parameter high and speakers will both incorporate new exemplars and deploy socially appropriate ones with alacrity; turn the sensitivity down low and speakers will be slower to learn new forms as well as less likely to respond to context.

Differential *LS* would affect the production of (o)/(oh) and (aw)T in the present data in the following way. Individuals with higher degrees of *LS* would be predicted to more readily acquire the (o)/(oh) distinction as a result of interaction with NYaE speakers, and to more readily exhibit Canadian Raising in the context of an interview that focuses on Canadian identity. Speakers with lower degrees of *LS* would lag in the acquisition of the low back vowel distinction, but simultaneously be less likely to activate their raised Canadian vowels in the interview setting. The dispersion of speakers shown in Fig 6.2, following a diagonal running from low degree of Canadian Raising and small distance between (o) and (oh) to high degree of Canadian Raising and high differentiation of low back vowels, thus follows from such an account.

At this point, however, the *linguistic sensitivity* account is mere hypothesis. Fortunately, it is testable in future research in several ways. The most direct way to test this account is to establish that the subjects do vary in conversational sensitivity, by administering something like the questionnaire of Likert-scale items described in Daly and Diesel (1992), and then establishing that this measure covaries with the sort of linguistic behavior described in this dissertation; in future expansions of this study, I plan to administer some modification of this questionnaire to interviewees. However, this proposal generates additional predictions which may be tested with existing data or data as yet to be collected.

For example, I suggested in the previous section that the high levels of raising shown by many of these speakers is a consequence of the interview situation, in which Canadian identity is made especially salient. This account implies that in a conversational context in which Canadian identity is not a salient feature, (aw)T raising ought to be attenuated, while low back vowel distance should remain essentially the same. The addition of LS to the model brings with it an additional prediction: the difference in (aw) performance across contexts should be greatest for the most sensitive speakers. This could be tested by, for example, by having native CE speakers read a series of passages containing many (aw) words, some of which focus on Canada-evoking topics, and others which are neutral with respect to markers of Canadian identity. While some shift towards more raised (aw)T realizations in the Canada-evoking context is predicted to occur for everyone, this shift should be greatest for speakers who are independently identified as more sensitive.

8.5 Implications for a theory of representation

The goal of this study was to use second dialect acquisition data to test the predictions made by two competing views of phonological representation. Here I review the major predictions made by each type of theory, and assess the extent to which they were borne out in this data.

A generative account of accommodation predicts relative stability with respect to low back vowel merger, as acquiring a new contrast requires changing underlying representations for many lexical items. This is to an extent borne out by the data, which do reflect a certain amount of first dialect stability. In minimal pair contexts, speakers remain merged in production and perception. In word list or conversational context, where most speakers do exhibit a distinction, the difference is rather small; they do not approach the robust distinction of many NYaE speakers. It is apparent then, that old representations have not disappeared entirely.

In contrast, the generative account allows for various simple changes to the Canadian Raising rule: the rule can be eliminated altogether, made to apply variably, or changed to result in a phonetically altered output. None of these predictions are borne out by the data here. The rule has clearly not been eliminated by the speakers in this study, as nearly all continue to exhibit Canadian Raising in some context. Moreover, this raising shows frequency effects which are not consistent with a single change in the generalization.

The word-specific effects noted for both the low back vowels and (aw) might be accommodated within a generative framework via the introduction of many lexically specific phonetic adjustment rules. At best this is an inelegant solution; at worst it starts to look rather similar to a theory of word-specific phonetics.

Finally, this account does not make correct predictions about the relative accommodatability of these two features. Given the formal simplicity of altering one rule versus altering potentially hundreds of individual lexical representations, the generative account implies that speakers will show signs of altering the Canadian Raising rule before they show signs of a low back vowel contrast. However, this relationship is not visible in the data. Instead, we find that speakers whose low back vowels diverge continue to exhibit Canadian Raising, which seems best explained in terms of their differing social significance.

A usage-based theory of accommodation predicts gradual divergence of low back vowel words, resulting in the eventual acquisition of a distinction between these word classes. This is what is seen in the data. The phonetic distance is small, and the token clouds for any given speaker show a high degree of overlap. Moreover, it is unclear whether speakers would be able to rely heavily on this difference in perception, for example to discriminate minimal pairs. This may be due to the actual exposure these speakers have had to the low back vowel contrast; it could be that many of them do not interact with speakers who have the classic NYAE distinction, but instead work or socialize with speakers who exhibit a subtler contrast. Or, it may indicate the strong influence of old D1 exemplars which have yet (if ever) to be overcome by the acquisition of D2 representations. This latter account is corroborated by the merged performance on the minimal pair tasks. When speakers are asked for judgments about how these words are pronounced in the minimal pair task, they are possibly accessing older representations more strongly associated with the self and representations acquired in childhood. However subtle the surface realization of this difference, is does exist and is systematic, which means that it needs to be represented.

Usage-based theory also predicts gradual convergence of pre-voiceless (aw) tokens with (aw) in other contexts. This is also borne out in the data, at least among less salient raising class words. Again, there is stability to the system, reflecting the influence of old exemplars. The results pertaining to *about* and *out* are puzzling in a purely linguistic usage-based theory, but when this theory incorporates social labels that indicate aspects of identity, they become more understandable: tokens of raised (aw)T are labeled *Canadian* and preferentially activated in contexts that evoke that identity, and this is especially true for shibboleth words that are strongly associated with that label and identity. This use of social labels, combined with the parameter of linguistic sensitivity introduced above, also provides an account of the positive correlation found between degree of low back vowel distance and Canadian Raising.

In summary, the data presented in this study indicates the need for a theory which incorporates both phonetically rich representations and social labels.

However, first dialect stability and a level of abstraction is still evident. Both of these aspects may be accommodated within the sort of hybrid model introduced in Section 3.3. Features of the first variety acquired in life may be stored in the abstract, 'cortical' component of such a hybrid model, while later changes made to the representation of word-level categories seem to be realized in the usage-based 'hippocampal' component.

8.6 Future research

Science progresses via a constant back-and-forth between data and theory: Observe some data, formulate an account, generate predictions, find data to test those predictions, refine the theory, derive new predictions; lather, rinse, repeat. This dialogue between the theoretical and the empirical has occurred throughout the course of this project. I began with two theories and the predictions each made about the acquisition of second dialect features, collected and analysed the relevant data in a way that these predictions could be tested, and arrived at a new theoretical account. Of course, this account brings with it new predictions that need to be tested, though doing so is beyond the scope of this dissertation. In the remainder of this section, I outline some of the questions raised by this research, and how future research by me or others might address them.

Establishing linguistic sensitivity and its effects. I have already outlined a few avenues for future research on this point in the sections above. First, *linguistic sensitivity* must be established as measurable and shown to correlate with the sort of linguistic behavior observed in this study. An obvious place to begin is with the measures of conversational sensitivity used by e.g. Daly and Diesel (1992), though these probably ought to be expanded to include sensitivity to socially meaningful linguistic forms. Once some measure of LS has been established, it can be tested as a predictor of other patterns of linguistic behavior, for instance (as described above) variations in magnitude of style shifting across contexts. If it can be shown that speaker sensitivity is related to the accommodation to and use of linguistic variables described here, then that will be compelling supporting evidence for this proposal.

In addition to predictions about style shifting, the LS account posits that speakers with high sensitivity are also those who have been the most successful at absorbing second dialect exemplars and thus at acquiring the low back vowel contrast. Again, this claim could be tested in future research, in a number of ways. One of these might be a slight addition to the experimental paradigm used by Goldinger (1998) and Nielsen (2006). These studies show that speakers make fine-grained adjustments in their productions of words based on previously heard target exemplars. Surely, not all of the subjects in these studies shift to the same extent; an obvious add-on to such research would be to collect imitative data from speakers, assess their sensitivity with suitable survey instruments, and then see whether interspeaker variation in the study results is at all explained by the variation in sensitivity. Those speakers with higher sensitivity should also be the ones to show the greatest degree of shift in the experimental condition.

Additional variables in this dataset. Only two linguistic variables were analyzed in the current study. Of course, the speakers produced many more word classes than (o), (oh) and (aw). Future analysis of this data will examine other variables which differ between CE and NYAE. One feature which seems apt for study is realization of (æ). Accommodation to this feature presents a somewhat different linguistic task for speakers of CE who are exposed to NYAE input: while both dialects share roughly similar allophones (tense [æ] and lax [æ]), the distribution of these allophones across phonological contexts is different. Studying this feature could thus shed light on how speakers re-allocate existing allophones to reflect new dialect input.

Different mobile populations. Another obvious extension of this work is to look at the ways in which different populations of mobile speakers acquire the features of their new dialect region. For example, a study of Canadians moving to another region which is also characterized by a low back vowel contrast and no raising, such as southern England, would provide an important replication of this study, while varying the precise social context: Southern England is associated with a prestige variety of English, in contrast to the 'great sink of negative prestige' that is New York City (Labov 1966).

It would also be informative to look at the behavior of a reciprocal population: native New Yorkers who move to Canada. Such speakers would be faced with the reverse linguistic situation of the speakers in this study: in order to accommodate, they would have to lose or at least attenuate their native (o)/(oh) contrast while acquiring Canadian Raising. The usage-based account advocated here also makes specific predictions about the manner in which these changes in an individual should proceed: phonetically and lexically gradual merger of the (o) and (oh) classes should occur, alongside gradual separation of (or more likely, enhancement of the existing phonetic separation between) (aw) found in raising versus nonraising contexts. Due to the differential social marking of these two features, (o)/(oh) merger is predicted to occur to a greater extent than adoption of Canadian Raising, with the latter feature showing more stylistic and context-dependent variation.

APPENDIX A
THE WORD LIST

marry	bait	tool	go	but	full
fool	pass	trap	pen	boot	dot
caller	mary	wok	start	logger	cat
fare	copy	whale	piece	foot	water
been	bought	bite	don	new	dad
pad	cute	odd	bad	bomb	face
sock	thought	cure	law	by	awed
boat	north	paw	tide	talk	pit
card	lager	bat	pull	coal	taught
boy	hurry	pan	bowl	feel	bag
la	sorry	palm	cot	dog	bet
dress	pond	merry	pin	man	walk
caught	mess	force	coffee	goat	fill
price	tab	pat	house	kit	knotty
stat	prince	collar	ferry	tall	tap
good	pawn	tour	tore	dew	cog
cow	pa	daughter	beet	doll	do
square	fleece	cut	nurse	dawn	pill
pot	higher	goose	cloth	fire	bade
near	mice	born	put	pawned	bit
furry	path	barn	bed	lot	sad
father	about	choice	bath	mouth	bother
naughty	pole				

APPENDIX B

MINIMAL AND RHYMING PAIR LIST

B.1 Minimal Pair List

prints prince	wail whale	Mary mary	coal call
Don Dawn	pull pole	tour tore	ferry furry
merry Mary	caught cot	knotty naughty	logger lager
bad bed	pawned pond	fill feel	odd awed
fool full	caller collar	pa paw	pen pin
la law	barn born	do dew	walk wok

B.2 Rhyming Pair List

pit	boot	cog	pen
kit	foot	dog	man
bought	fire	father	doll
pot	higher	bother	tall
bad	palm	daughter	talk
dad	bomb	water	sock

BIBLIOGRAPHY

Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. The University of Chicago Press.

Anttila, Arto. 1997. Variation in Finnish phonology and morphology. Doctoral Dissertation, Stanford University.

Anttila, Arto. 2002a. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20:1–42.

Anttila, Arto. 2002b. Variation and phonological theory. In *The handbook of language variation and change*, ed. Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes, 206–243. Blackwell.

Ash, Sharon. 2002. The distribution of a phonemic split in the Mid-Atlantic region: Yet more on short a. *Penn Working Papers in Linguistics* 8.3:1–15.

Auger, Julie. 2001. Phonological variation and Optimality Theory: Evidence from word-initial vowel epenthesis in Vimeu Picard. *Language Variation and Change* 13:253–303.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics*. Cambridge University Press.

Baayen, R. Harald, Richard Piepenbrock, and Hed derik van Rijn. 1993. *The CELEX lexical database*. Linguistic Data Consortium, University of Pennsylvania.

Bakker, Arnold B., Karen I. Van der Zee, Kerry A. Lewig, and Maureen F. Dollard. 2006. The relationship between the big five personality factors and burnout: A study among volunteer counselors. *The Journal of Social Psychology* 146:31–50.

Bakovic, Eric, and Bozene Pajak. 2008. Contingent optionality. Paper presented at the 82nd Linguistic Society of American meeting, Chicago.

Bates, Douglas, and Deepayan Sarkar. 2008. lme4: linear mixed-effects models using s4 classes. http://cran.r-project.org.

Bates, Elizabeth, Inge Bretherton, and Lynn Snyder. 1988. *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge University Press.

Becker, Kara. 2010. Social conflict and social practice on the Lower East Side: A study of regional dialect features in New York City English. Doctoral Dissertation, New York University.

Becker, Kara, and Amy W. Wong. 2009. The short-a system of NYC English: An update. *Penn Working Papers in Linguistics* 14.

Bell, Allan. 1984. Language style a audience design. *Language in Society* 13:145–204.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14:150–177.

Bernieri, Frank J., and Robert Rosenthal. 1991. Interpersonal coordination: Behavior matching and interactional synchrony. In *Fundamentals of non-verbal behavior*, ed. Robert S. Feldman and Bernard Rime, 401–432. Cambridge University Press.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. Winifred Strange, 171–204. York Press.

Boberg, Charles. 2000. Geolinguistic diffusion and the U.S.-Canada border. *Language Variation and Change* 12:1–24.

Boberg, Charles. 2005. The Canadian Shift in Montreal. *Language Variation and Change* 17:133–154.

Boberg, Charles. 2008. English in Canada: Phonology. In *Varieties of English: The Americas and the Caribbean*, ed. Edgar W. Schneider, volume 2, 144–160. Mouton de Gruyter.

Boberg, Charles. 2009. The emergence of a new phoneme: Foreign (a) in Canadian English. *Language Variation and Change* 21:355–380.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.

Bowie, David. 2000. The effect of geographic mobility on the retention of a local dialect. Doctoral Dissertation, University of Pennsylvania.

Bowie, David. 2009. The ageing voice: Changing identity over time. In *Language and identities*, ed. Carmen Llamas and Dominic Watt, 55–66. Edinburgh University Press.

Brown, Vivian. 1990. The social and linguistic history of a merger: /i/ and /e/ before nasals in Southern American English. Doctoral Dissertation, Texas A and M.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge University Press.

Cappella, Joseph N. 1981. Mutual influence in expressive behavior: Adult-adult and infant-adult dyadic interaction. *Psychological Bulletin* 89:101–132.

Cappella, Joseph N. 1996. Dynamic coordination of vocal and kinesic behavior in dyadic interaction: Methods, problems, and interpersonal outcomes. In *Dynamic patterns in communication processes*, ed. James H. Watt and C. Arthur VanLear, 353–386. Sage Publications.

Cardoso, Walcir. 2001. Variation patterns in across-word regressive assimilation in Picard: An Optimality Theoretic account. *Language Variation and Change* 13:305–341.

Cedergren, Henrietta J., and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50:333–355.

Chambers, Jack K. 1973. Canadian Raising. *Canadian Journal of Linguistics* 18:113–135.

Chambers, Jack K. 1988. Acquisition of phonological variants. In *Methods in dialectology*, ed. Alan R. Thomas. Multilingual Matters.

Chambers, Jack K. 1992. Dialect acquisition. *Language* 68:673–705.

Chambers, Jack K. 2002. Dynamics of dialect convergence. *Journal of Sociolinguistics* 6:117–130.

Chambers, Jack K., and Peter Trudgill. 1998. *Dialectology*. Cambridge University Press, 2nd edition.

Chartrand, Tanya L., and John A. Bargh. 1999. The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology* 76:893–910.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Harper and Row.

Church, Barbara A., and Daniel L. Schacter. 1994. Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20:521–533.

Clarke, Sandra, Ford Elms, and Amani Youssef. 1995. The third dialect of English: Some Canadian evidence. *Language Variation and Change* 7:209–228.

Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2:225–252.

Clements, George N., and Elizabeth Hume. 1995. The internal organization of speech sounds. In *The handbook of phonological theory*, ed. John A. Goldsmith. Blackwell.

Coetzee, Andries. 2004. What it means to be a loser: Non-optimal candidates in Optimality Theory. Doctoral Dissertation, University of Massachusetts Amhert.

Coetzee, Andries. 2006. Variation as accessing "non-optimal" candidates. *Phonology* 23:337–385.

Coetzee, Andries. 2009. Phonological variation and lexical frequency. In *NELS 38*, ed. Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, volume 1, 189–202.

Coetzee, Andries, and Joe Pater. To Appear. The place of variation in phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu. Blackwell, 2nd edition.

Cohen, Paul. 1970. The tensing and raising of short [a] in the metropolitan area of New York City. Master's thesis, Columbia University.

Cole, Ronald A., Max Coltheart, and Fran Allard. 1974. Memory of a speaker's voice: Reaction time to same- or different-voiced letters. *Quarterly Journal of Experimental Psychology* 26:1–7.

Condon, William S., and Louis W. Sander. 1974. Neonate movement is synchronized with adult speech: Interactional participation and language acquisition. *Science* 183:99–101.

Conn, Jeffrey, and Uri Horesh. 2002. Assessing the acquisition of dialect variables by migrant adults in Philadelphia. *Penn Working Papers in Linguistics* 8.3.

Coupland, Nikolas. 1984. Accommodation at work. *International Journal of the Sociology of Language* 4-6:49–70.

Coye, Dale F. 2009. Dialect boundaries in New Jersey. *American Speech* 84:414–452.

Daly, John A., and Carol A. Diesel. 1992. Measures of communication-related personality variables. *Communication Education* 41:405–415.

Daly, John A., Anita L. Vangelisti, and Suzanne M. Daughton. 1987. The nature and correlates of conversational sensitivity. *Human Communication Research* 14:167–202.

De Decker, Paul. 2006. Migration: social mobility and phonetic drift. *Penn Working Papers in Linguistics: Selected paper from NWAV 34* .

De Decker, Paul, and Jennifer Nycz. 2006. Are tense [æ]s really tense? an ultrasound study. *Penn Working Papers in Linguistics: Proceedings of the 29th Annual Penn Linguistics Colloquium* .

DeCamp, David. 1972. Hypercorrection and rule generalization. *Language in Society* 1:87–90.

Dell, Francois. 1981. On the learnability of optional phonological rules. *Linguistic Inquiry* 12:31–37.

Delvaux, Veronique, and Alain Soquet. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64:145–173.

Di Paolo, Marianna. 1988. Pronunciation and categorization in sound change. In *Linguistic change and contact: NWAV-XVI*, ed. Kathleen Ferrara, Becky Brown, Keith Walters, and John Baugh, 84–92. Department of Linguistics, University of Texas.

Dinnsen, Daniel A., and Jan Charles-Luce. 1984. Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics* 12:49–60.

Drager, Katie. 2009. A sociophonetic ethnography of Selwyn Girls' High. Doctoral Dissertation, University of Canterbury.

Dresher, B. Elan. 1993. The chronology and status of Anglian smoothing. In *Phonetics and phonology: Studies in lexical phonology*, ed. Sharon Hargus and Ellen M. Kaisse, 325–342. Academic Press.

Eckert, Penelope. 2000. *Linguistic variation as social practice*. Blackwell.

Eisner, Frank, and James M. McQueen. 2006. Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America* 119:1950–1953.

Evans, Bronwen G., and Paul Iverson. 2007. Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America* 121:3814–3826.

Fant, Gunnar. 1960. *Acoustic theory of speech production*. The Hague.

Flege, James E. 1987. The production of '“new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15:47–65.

Flege, James E., and James Hillenbrand. 1986. Differential use of temporal cues to the [s-z] contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America* 79:508–517.

Foulkes, Paul. 2010. Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology* 1:5–39.

Foulkes, Paul, and Gerard Docherty. 2006. The social life of phonetics and phonology. *Journal of Phonetics* 34:409–438.

Giannini, Antonella, and Umberto Cinque. 1978. Phonetic status and phonemic function of the final devoiced stops in Polish. *Speech Laboratory Report* .

Giles, Howard. 1973. Accent mobility: A model and some data. *Anthropological Linguistics* 15:87–105.

Giles, Howard, Justine Coupland, and Nikolas Coupland. 1991. Accommodation theory: Communication, context, and consequence. In *Contexts of accommodation: Developments in applied sociolinguistics*, ed. Howard Giles, Justine Coupland, and Nikolas Coupland, 1–68. Cambridge University Press.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105:251–279.

Goldinger, Stephen D. 2007. A complementary systems approach to abstract and episodic speech perception. *ICPhS-XVI* 49–54.

Goldinger, Stephen D., and Tamiko Azuma. 2004. Episodic memory reflected in printed word naming. *Psychonomic Bulletin and Review* 11:716–722.

Goldsmith, John A. 1979. The aims of autosegmental phonology. In *Current approaches to phonological theory*, ed. D.A. Dinnsen, 202–222. Indiana University Press.

Goldsmith, John A. 1995. Phonological theory. In *The handbook of phonological theory*, ed. John A. Goldsmith, 1–23. Blackwell.

Gordon, Matthew. 2001. Missouri is for mergers. Paper presented at NWAVE 30.

Gore, Jonathan S., and Mary Jill Rogers. 2010. Why do I study? The moderating effect of attachment style on academic motivation. *The Journal of Social Psychology* 150:560–578.

Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. In *CLS 35*, 151–66. University of Chicago.

Gregory, Stanford W. 1990. Analysis of fundamental frequency reveals covariation in interview partner's speech. *Journal of Nonverbal Behavior* 14:237–251.

Guy, Gregory R. 1991a. Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3:223–239.

Guy, Gregory R. 1991b. Explanation in a variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3:1–22.

Guy, Gregory R. 1997. Violable is variable: Optimality Theory and linguistic variation. *Language Variation and Change* 9:333–347.

Guy, Gregory R. 2007. Variation and phonological theory. In *Sociolinguistic variation: Theories, methods and applications*, ed. Robert Bayley and Ceil Lucas, 5–23. Cambridge University Press.

Guy, Gregory R., and Charles Boberg. 1997. Inherent variability and the obligatory contour principle. *Language Variation and Change* 9:149–164.

Halle, Morris. 1985. Speculations about the representations of words in memory. In *Phonetic linguistics*, ed. Victoria Fromkin, 101–114. Academic Press.

Harrington, Jonathan. 2006. An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts. *Journal of Phonetics* 34:439–457.

Harrington, Jonathan, SallyAnne Palethorpe, and Catherine Watson. 2000. Monophthongal vowel changes in recieved pronunciation: An acoustic analysis of the queen's christmas broadcasts. *Journal of the International Phonetic Association* 30:63–78.

Harris, Zellig. 1942. Morpheme alternants in linguistic analysis. *Language* 18:169–180.

Hay, Jennifer, and Katie Drager. 2007. Sociophonetics. *Annual Review of Anthropology* 36:89–103.

Hay, Jennifer, Aaron Nolan, and Katie Drager. 2006a. From fush to feesh: Exemplar priming in speech perception. *Linguistic Review* 23:351–379.

Hay, Jennifer, Paul Warren, and Katie Drager. 2006b. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34:458–484.

Hazen, Kirk. 2003. Division and unity in West Virginia mergers. Paper presented at NWAVE 32.

Herold, Ruth. 1990. Mechanisms of merger: The implementation and distribution of the low back merger in Pennsylvania. Doctoral Dissertation, University of Pennsylvania.

Herzog, Marvin. 1965. *The Yiddish language in northern Poland*. Bloomington and The Hague.

Hintzman, Douglas L. 1986. "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93:411–428.

Hintzman, Douglas L., Richard A. Block, and Norman R. Inskeep. 1972. Memory for mode of input. *Journal of Verbal Learning and Verbal Behavior* 11:741–749.

Irons, Terry. 2005. An experimental evaluation of a possible mechanism for back vowel merger. Paper presented at NWAVE 34.

Jaeger, T. Florian, and Laura Staum. 2005. That-omission beyond processing: stylistic and social effects. *Paper presented at NWAV 34, Ohio State.* .

Jakobson, Roman. 1962. *Selected writings, vol. 1*. Mouton and Company.

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

Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3:359–383.

Johnson, Keith. 1997. Speech perception without speaker normalization. In *Talker variability in speech processing*, ed. Keith Johnson and John W. Mullennix, 145–166. Academic Press.

Johnson, Keith. 2006. Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34:485–499.

Joos, Martin. 1942. A phonological dilemma in Canadian English. *Language* 18:141–144.

Joos, Martin. 1948. Acoustic phonetics. *Language* 24.

Jurafsky, Daniel, Alan Bell, Eric Fosler-Lussier, Cynthia Girand, and William D. Raymond. 1998. Reduction of English function words in Switchboard. In *ICSLP-98*, volume 7, 3111–3114. Sydney.

Kaisse, Ellen M. 1993. Rule reordering and rule generalization in lexical phonology: A reconsideration. In *Phonetics and phonology: Studies in lexical phonology*, ed. Sharon Hargus and Ellen M. Kaisse, 343–365. Academic Press.

Keating, Patricia. 1988. Underspecification in phonetics. *Phonology* 5:275–292.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Blackwell.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in the morning calm*, ed. In-Seok Yang, volume 2, 3–91. Seoul: Hanshin.

Kiparsky, Paul. 1983. Word formation and the lexicon. In *Proceedings of the Mid-America Linguistics Conference*, ed. Frances Ingemann, 3–29. Lawrence: University of Kansas.

Kiparsky, Paul. 1988. Phonological change. In *Linguistics: The Cambridge survey*, ed. Frederick J. Newmeyer, volume 1, 363–415. Cambridge University Press.

Kirov, Christo, and Adamantios I. Gafos. 2007. Dynamic phonetic detail in lexical representations. In *ICPhS XVI*, 637–640.

Kroch, Anthony. 1994. Morphosyntactic variation. In *Papers from the 30th regional meeting of the Chicago Linguistic Society vol. 2: The parasession on variation* *in linguistic theory*, ed. Katherine Beals, Jeannette Denton, Robert Knippen, Lynette Melnar, Hisam Suzuki, and Erica Zeinfeld, 180–201. Chicago Linguistic Society.

Kučera, Henry, and W.Nelson Francis. 1967. *Computational analysis of present-day American English*. Brown University Press.

Kuhl, Patricia K., and Andrew N. Meltzoff. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America* 100:2425–2438.

Kurath, Hans, and Raven I. McDavid. 1961. *The pronunciation of English in the atlantic states*. Ann Arbor.

Labov, William. 1963. The social motivation of a sound change. *Word* 19:273–309.

Labov, William. 1966. *The social stratification of English in New York City*. Center for Applied Linguistics, Washington D.C., 1st edition.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762.

Labov, William. 1972. *Sociolinguistic patterns*. University of Pennsylvania Press.

Labov, William. 1991. The three dialects of English. In *New ways of analyzing sound change*, ed. Penelope Eckert, 1–44. Academic Press.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Blackwell.

Labov, William. 2006a. *The social stratification of English in New York City*. Cambridge University Press, 2nd edition.

Labov, William. 2006b. A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics* 34:500–515.

Labov, William, Sharon Ash, and Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology, and sound change: A multimedia reference tool*. Walter de Gruyter.

Labov, William, Malcah Yaeger, and Richard Steiner. 1972. *A quantitative study of sound change in progress*. U.S. Regional Survey.

LaFrance, Marianne. 1982. Posture mirroring and rapport. In *Interaction rhythms: Periodicity in communicative behavior*, 279–298. Human Sciences Press.

Langacker, Ronald. 1987. *Foundations of cognitive grammar, volume 1: Theoretical perspectives*. Stanford University Press.

Langacker, Ronald. 2000. A dynamic usage-based model. In *Usage-based models of language*, ed. Michael Barlow and Susanne Kemmer, 1–63. CSLI.

Lenneberg, Eric. 1967. *Biological foundations of language*. Wiley.

Majors, Tivoli. 2005. Low back vowel merger in Missouri speech: Acoustic description and explanation. *American Speech* 80:165–179.

Malhotra, Naresh K., Arun K. Jain, and Christian Pinson. 1983. Extremity of judgment and personality variables: Two empirical investigations. *The Journal of Social Psychology* 120:111–118.

McClelland, James L., and Jeffrey L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18:1–86.

Mendoza-Denton, Norma, Jennifer Hay, and Stefanie Jannedy. 2003. Probabilistic sociolinguistics: Beyond variable rules. In *Probabilistic linguistics*, ed. Rens Bod, Jennifer Hay, and Stefanie Jannedy, 97–138. MIT Press.

Milroy, Lesley. 1987. *Language and social networks*. Wiley-Blackwell, 2nd edition.

Montgomery, Michael. 1994. The evolution of verb concord in scots. In *Studies in Scots and Gaelic: Proceedings of the Third International Conference on the Languages of Scotland*, ed. Alexander Fenton and Donald A. MacDonald. Cannongate Academic.

Morton, John. 1969. Interaction of information in word recognition. *Psychological Review* 76:165–178.

Mullennix, John W., David B. Pisoni, and Christopher S. Martin. 1988. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America* 85:365–378.

Munro, Murray J., Tracey M. Derwing, and James E. Flege. 1999. Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics* 27:385–403.

Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe, and Aubrey R. White. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics* 34:202–240.

Myers, James. 1994. Rules, constraints, and lexical phonology in Glenoe Scots. Ms., December 1994.

Nagy, Naomi, and William T. Reynolds. 1997. Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change* 9:37–55.

Nahkola, Kari, and Marja Saanilahti. 2004. Mapping language changes in real time: A panel study on Finnish. *Language Variation and Change* 16:75–92.

Natale, Michael. 1975. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology* 32:790–804.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18:62–85.

Nielsen, Kuniko. 2005. Generalization of phonetic imitation across place of articulation. In *Proceedings for ISCA Workshop on Plasticity in Speech Perception*. University College London, London, UK.

Nielsen, Kuniko. 2006. VOT imitation generalizes across phonemes and natural classes. Paper presented at the 2006 LSA Meeting in Albuquerque, NM.

Norris, Dennis, James McQueen, and Anne Cutler. 2003. Perceptual learning in speech. *Cognitive Psychology* 47:204–238.

Nycz, Jennifer. 2005. The dynamics of near-merger in accommodation. In *Proceedings of ConSOLE XIII*, ed. Sylvia Blaho, Luis Vicente, and Erik Schoorlemmer. http://www.hum.leiden.edu/lucl/research/sole/proceedings/console13.html.

Nycz, Jennifer. 2006. Global faithfulness and choice of repair. *Penn Working Papers in Linguistics* 10:205–218.

Nygaard, Lynne C., and David B. Pisoni. 1998. Talker-specific learning in speech perception. *Perception and Psychophysics* 60:355–376.

O'Dell, Michael, and Robert Port. 1983. Discrimination of word-final voicing in German. Paper presented at the Acoustical Society of America, Cincinnati.

Office of Management and Budget. 2009. Update of statistical area definitions and guidance on their uses. OMB Bulletin No. 10-02.

van Oostendorp, Marc. 1997. Style levels in conflict resolution. In *Variation, change, and phonological theory*, ed. Frans Hinskens, Roeland van Hout, and W. Leo Wetzels, 207–229. Amsterdam: John Benjamins.

Pardo, Jennifer. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119:2382–2393.

Pater, Joseph. 2009. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In *Phonological argumentation: Essays on evidence and motivation*, ed. Steve Parker, 123–154. Equinox.

Paul, Hermann. 1889. Principles of the history of language. trans. H.A. Strong.

Payne, Arvilla. 1976. The acquisition of the phonological system of a second dialect. Doctoral Dissertation, University of Pennsylvania.

Phillips, Betty. 1984. Word frequency and the actuation of sound change. *Language* 60:320–342.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In *Frequency and the emergence of linguistic structure*, ed. Joan Bybee and Paul J. Hopper, 137–157. John Benjamins.

Pierrehumbert, Janet. 2002. Word-specific phonetics. In *Laboratory Phonology 7*, ed. Carlos Gussenhoven and Natasha Warner. Mouton de Gruyter.

Pierrehumbert, Janet. 2003. Probabilistic phonology: Discrimination and robustness. In *Probabilistic linguistics*, ed. Rens Bod, Jennifer Hay, and Stefanie Jannedy. MIT Press.

Pierrehumbert, Janet. 2006. The next toolkit. *Journal of Phonetics* 34:516–530.

Pierrehumbert, Janet, Mary Beckman, and D. R. Ladd. 2000. Conceptual foundations of phonology as a laboratory science. In *Phonological knowledge: Conceptual and empirical issues*, ed. Noel Burton-Roberts, Philip Carr, and Gerard Docherty, 273–303. Oxford University Press.

Pinheiro, Jose C., and Douglas M. Bates. 2000. *Mixed-effect models in S and S-Plus*. Springer.

Port, Robert, and Penny Crawford. 1989. Incomplete neutralization and pragmatics in German. *Journal of Phonetics* 17:2257–2282.

Port, Robert, Fares Mitleb, and Michael O'Dell. 1981. Neutralization of obstruent voicing in German is incomplete. Paper presented at the Acoustical Society of America, Miami.

Preston, Dennis. 1999. A language attitude approach to the perception of regional variety. In *Handbook of perceptual dialectology*, ed. Dennis Preston, volume 1. John Benjamins.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Technical Report 2, Rutgers University.

Reynolds, Bill. 1994. Variation and phonological theory. Doctoral Dissertation, University of Pennsylvania.

Sancier, Michele L., and Carol A. Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25:421–436.

Sankoff, David, and William Labov. 1979. On the uses of variable rules. *Language in Society* 8:189–222.

Sankoff, Gillian. 2004. Adolescents, young adults, and the critical period: Two case studies from 'Seven Up'. In *Sociolinguistic variation: Critical reflections*, ed. Carmen Fought, 121–139. Oxford University Press.

Sankoff, Gillian, and Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83:560–588.

de Saussure, Ferdinand. 1916. *Cours de linguistique générale*. Payot.

Schacter, Daniel L., and Barbara A. Church. 1992. Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18:915–930.

Schilling-Estes, Natalie. 1998. Investigating 'self-concious' speech: The performance register in Ocracoke English. *Language in Society* 27:53–83.

Scobbie, James M., and Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In *Contrast in phonology: Theory, perception, acquisition*, ed. Peter Avery, B. Elan Dresher, and Keren Rice, 87–113. Mouton de Gruyter.

Shockey, Linda. 1984. All in a flap: long-term accommodation in phonology. *International Journal of the Sociology of Language* 46:87–95.

Shockley, Kevin, Laura Sabadini, and Carol A. Fowler. 2004. Imitation in shadowing words. *Perception and Psychophysics* 66:422–429.

Shockley, Kevin, Marie-Vee Santana, and Carol A. Fowler. 2003. Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance* 29:326–332.

Smolensky, Paul. 1996. The initial state and 'richness of the base' in Optimality Theory. Technical report, Johns Hopkins University.

Soderstrom, Melanie, Donald Mathis, and Paul Smolensky. 2006. Abstract genomic coding of Universal Grammar in Optimality Theory. In *The harmonic mind*, 403–471. Paul Smolensky and Geraldine Legendre.

Studdert-Kennedy, Michael. 1976. Speech perception. In *Contemporary issues in experimental phonetics*, ed. Norman J. Lass. Academic Press.

Tagliamonte, Sali A. 1998. Was/were variation across the generations: View from the city of York. *Language Variation and Change* 10:153–91.

Tagliamonte, Sali A., and Sonja Molfenter. 2007. How'd you get that accent?: Acquiring a second dialect of the same language. *Language in Society* 36:649–675.

Tesar, Bruce, and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29:229–268.

Tillery, Jan, and Guy Bailey. 2003. Approaches to real time in dialectology and sociolinguistics. *World Englishes* 22:351–365.

Trubetzkoy, Nikolai Sergeyevich. 1969[1939]. Principles of phonology. trans. Christiane A. M. Baltaxe.

Trudgill, Peter. 1974. *The social differentiation of English in Norwich*. Cambridge University Press.

Trudgill, Peter, and Nina Foxcroft. 1978. On the sociolinguistics of vocalic mergers: Transfer and approximation in East Anglia. In *Sociolinguistic patterns in british english*, ed. Peter Trudgill, 69–79. Edwin Arnold.

Vaux, Bert. 2003. Syllabification in Armenian, universal grammar, and the lexicon. *Linguistic Inquiry* 34:91–125.

Vaux, Bert. 2008. Why the phonological component must be serial and rule-based. In *Rules, constraints and phonological phenomena*, ed. Bert Vaux and Andrew Nevins. Oxford University Press.

Venables, William N., David M. Smith, and the R Development Core Team. 2010. *An introduction to R (notes on R: A programming environment for data analysis and graphics)*. http://www.r-project.org/, Version 2.11.1.

Wagner, Suzanne Evans. 2008. Linguistic change and stabilization in the transition from adolescence to adulthood. Doctoral Dissertation, University of Pennsylvania.

Watt, Dominic, Carmen Llamas, and Daniel Ezra Johnson. 2010. Levels of linguistic accommodation across a national border. *Journal of English Linguistics* 38:270–289.

Wedel, Andrew. 2004. Category competition drives contrast maintenance within an exemplar-based production/perception loop. In *Proceedings of the Seventh*

Meeting of the ACL Special Interest Group in Computational Phonology, ed. John Goldsmith and Richard Wicentowski, volume 7, 1–10. ACL.

Wedel, Andrew. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23:247–274.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for historical linguistics*, ed. Winfred Lehmann and Yakov Malkiel, 97–195. University of Texas Press.

Wells, John C. 1982. *Accents of English*. Cambridge University Press.

Yu, Alan C. L. 2010. Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PLOS One* 5.

Zajonc, R. B., Pamela K. Adelmann, Sheila T. Murphy, and Paula M. Niedenthal. 1987. Convergence in the physical appearance of spouses. *Motivation and Emotion* 11:335–346.

Zec, Draga. 1993. Rule domains and phonological change. In *Phonetics and phonology: Studies in lexical phonology*, ed. Sharon Hargus and Ellen M. Kaisse, 365–405. Academic Press.

Zubritskaya, Katya. 1997. Mechanism of sound change in optimality theory. *Language Variation and Change* 9:121–148.