First Language Phonetic Drift During Second Language Acquisition

by

Charles Bond Chang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Linguistics

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:
Professor Keith Johnson, Chair
Professor Susanne Gahl
Professor Sharon Inkelas
Professor Carla Hudson Kam

Fall 2010

The dissertation of Charles Bond Chang is approved:

Chair ________________ Date 10/7/2010

________________ Date 10/7/2010

________________ Date 10/7/2010

________________ Date 10/1/2010

University of California, Berkeley

Fall 2010

First Language Phonetic Drift During Second Language Acquisition

Abstract

First Language Phonetic Drift During Second Language Acquisition

by

Charles Bond Chang

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Keith Johnson, Chair

Despite abundant evidence of malleability in speech production, previous studies of the effects of late second-language learning on first-language production have been limited to advanced learners. This dissertation examines these effects in novice learners, finding that experience in a second language rapidly, and possibly inexorably, affects production of the native language. In a longitudinal study of Korean acquisition, native English-speaking adult learners ($n = 19$) produced the same English words at weekly intervals over the course of intensive elementary Korean classes. Results of two acoustic case studies indicate that experience with Korean rapidly influences the production of English, and that the effect is one of assimilation to phonetic properties of Korean. In case study 1, experience with Korean stop types is found to influence the production of English stop types in terms of voice onset time (VOT) and/or fundamental frequency (f_0) onset as early as the second week of Korean classes, resulting in the lengthening of VOT in English voiceless stops (in approximation to the longer VOT of the perceptually similar Korean aspirated stops) and the raising of f_0 onset following English voiced and voiceless stops (in approximation to the higher f_0 levels of Korean). Similarly, in case study 2, experience with the Korean vowel space is found to have a significant effect on production of the English vowel space, resulting in a general raising of females' English vowels in approximation to the overall higher Korean vowel space. These rapid effects of second-language experience on first-language production suggest that cross-language linkages are established from the onset of second-language learning, that they occur at multiple levels, and that they are based not on orthographic equivalence, but on phonetic and/or phonological proximity between languages. The findings are discussed with respect to current notions of cross-linguistic similarity, exemplar models of phonology, and language teaching and research practices.

To *Appa*

Contents

List of Figures

List of Tables

Acknowledgments

Although in practical terms this dissertation was completed during the last two years, it got its start long ago when, as a wide-eyed freshman in college, I fell hopelessly in love with the field of linguistics. Cupid was played primarily by Bert Vaux, who has a way of teaching that makes one wonder how anyone could decide *not* to major in linguistics. I have learned a great deal from Bert, as well as from a number of other excellent linguists over the course of the last eleven years: Andrew Garrett, Brechtje Post, Ian Maddieson, Jie Zhang, John Williams, Larry Hyman, Leanne Hinton, Line Mikkelsen, Lisa Lavoie, and Patrick Taylor, to name only a few. I am grateful for all the knowledge and analytical skills they have instilled in me with boundless energy, wit, and good humor.

I have benefited intellectually from running into the paths of many great teachers, but during the last two years in particular I have taken away the most from the teaching and advising of my thesis committee: Sharon Inkelas, Susanne Gahl, Carla Hudson Kam, and Keith Johnson. Collectively they have had the greatest impact on the shape and content of this dissertation, providing numerous suggestions that have improved this work immeasurably. Sharon is probably the clearest instructor I have ever learned from and offered many insightful comments on earlier versions of this work. Her advanced phonology course is arguably the best in the world and is what kept me interested in phonology even as I was becoming a phonetician. Susanne taught me not only in graduate school, but in college as well, and did much to train me early on in how to both think and write clearly about language; thankfully, her meticulous attention to detail has left its mark on this dissertation. Quite the opposite of what one thinks of as an "outside" committee member, Carla offered some of the quickest and most helpful feedback on drafts, even despite her busy schedule attending to a much more demanding student! Last but not least, Keith has been a wonderful thesis advisor—knowledgeable, encouraging, and kind—and has guided this research with just the right amount of input from the moment I decided what kind of study I wanted to do two years ago. When I think about what kind of scholar I want to be, it is the intelligence, acumen, and broad vision of my committee members that come first to mind. Thanks to them for making this the best dissertation it could be!

As a study of speech production, this dissertation is based upon some very time-consuming acoustic analysis work, which I am fortunate to have had help with at various points in time. Daiana Chang and Kevin Sitek both provided skilled assistance with substantial parts of the analysis, getting an enormous amount done in a careful and timely manner. I am tremendously appreciative of all their hard work and look forward to the great things they will accomplish in the future (with or without Praat).

The phonology lab at Berkeley is where many of the ideas in this dissertation were developed or presented, and in addition to the Berkeley "P" crowd already mentioned, I thank Anne Pycha, Christian DiCanio, Clara Cohen, Dominic Yu, Eurie Shin, Grant McGuire, John Ohala, John Sylak, Marc Ettlinger, Melinda Woodley, Molly

Babel, Reiko Kataoka, Ronald Sprouse, Russell Rhodes, Sam Tilsen, Shinae Kang, Shira Katseff, Will Chang, Yao Yao, and Yuni Kim for making the lab an incredibly intellectually stimulating place. Portions of this work were also presented at Yonsei University and meetings of the Linguistic Society of America and the Acoustical Society of America, and I have received helpful feedback at all of these venues.

Many more people deserve thanks for the pivotal roles they played in simply making this study possible. I am indebted to Jai-Ok Shim and the Fulbright Korean-American Educational Commission, as well as the Korean Language and Culture Center at Korea University, for helping me to conduct a study of this size with minimal administrative obstacles—진심으로 감사드립니다! In particular, Aimee Jachym, Alexis Stratton, Christine Arrozal, Jamal Grimes, Jessica Lee, and Luke Dolge were accommodating and thoughtful in matters both logistical and non-logistical. They drummed up participation; negotiated with the powers-that-be to reserve space for me; worked my experiments into an already packed schedule; provided batteries in dire emergencies; the list goes on and on. Without their unflagging support, this study could not have been carried to completion. This study would have been impossible, moreover, without the generous funding of the U.S. Department of Education, the National Science Foundation, the Berkeley Center for Korean Studies, and the Berkeley Department of Linguistics. I wholeheartedly thank these bodies for their financial support, as well as Belén Flores, Natalie Babler, and Paula Floro for their administrative expertise, which saved my academic life on more than one occasion.

The study participants themselves earn special praise for being enthusiastic and committed for a long period of time. Getting people who are not undergraduate students to do a linguistics experiment is not easy; getting them to do multiple experiments over the course of six weeks is another matter altogether. Add to this the extremely hot and humid weather of summertime in Korea, and you begin to realize how dedicated the learner participants had to be to complete this study in its entirety. If I could thank them by name, I would! I am deeply grateful to both them and the native speaker participants who provided control data.

Finally, I cannot overestimate just how important my friends and family have been in helping me get to the point of at last filing this dissertation. Thanks to my cohort, including Ange Strom-Weber, Jisup Hong, Michael Houser, and Zhenya Antić, for always making class something to look forward to; to Erin Haynes, Kanako Ito, and Russell Rhodes for living with me and living to tell about it; to Albert Ching, Dan Chen, Iksoo Kwon, Maziar Toosarvandani, Rekia Jina Jibrin, Veronica Rivera, Yao Yao, and Yuni Kim for being the best partners-in-crime; and to all the Berkeley students I have been lucky enough to teach, whose bright faces never failed to infuse me with energy on days when it was sorely needed. My family in Korea kept me going through oppressively hot summers of data collection with delicious food, air-conditioned office space, and an endless supply of hilarious family stories, and my family in the States have always filled my time at home with love and laughter. Without a doubt, they deserve the biggest thanks of all. *Appa*, this one is for you.

Chapter 1

Introduction

1.1 Significance of the Study

When we learn a second language in our adult years, what happens to how we pronounce our native language, the language we learned first in childhood? According to common wisdom, our own individual version of our native language is very much a part of who we are as adults and remains essentially the same as we grow older. But is this really the case? It has often been observed that when people come to live in a place where a different dialect of their language is spoken (such as when teenagers leave their hometowns to attend college in distant parts of the country), they return home sounding perceptibly different from when they left. There are clear social explanations for this sort of "accent shift" between dialects. However, it has also been observed that when people leave their country to live abroad and are immersed in a totally unrelated language for an extended period of time (such as in the Peace Corps or various study-abroad programs), they too return home sounding perceptibly different from when they left. Why should learning a foreign language affect an individual's native language production? How and when does this cross-language influence manifest itself? And what does this sort of cross-language influence reveal about how language sounds—of the native language and of a foreign language—are represented in the mind? These are the questions behind the line of research pursued in this dissertation.

This dissertation is a study of the phonetics of language contact. In particular, it is an experimental investigation of the phonetic developments that occur when two phonological systems come into contact within mature speakers of one language who are acquiring a second language. Insofar as the population of interest comprises second-language acquirers, this is a study of second language acquisition. However, in contrast to most research in second language acquisition, which investigates the influence of the first, or native, language system on the emerging second, or target, language system, the present study focuses on the reverse type of cross-language influence: the effect of the second-language system on the first-language system. In this respect, this dissertation is a study of first language development—specifically, of phonological restructuring in the first language during second language acquisition, a phenomenon that will be referred to here as PHONETIC DRIFT.

The study of phonetic drift, and of cross-linguistic influence more generally, holds far-reaching implications for both linguistic theory and practice. Investigations of phonetic drift shed light on two basic components of any theory of language: representation and development. By virtue of examining how native language production can be affected by disparate input, such studies elucidate the fundamental nature of the linguistic representations drawn upon in production, as well as the ways in which they may be connected to each other across languages. In addition, by recording the progression of native language adaptation to foreign input over time, this research broadens our understanding of language development—in particular, of the manner in which first-language structures acquired in childhood may continue to be tuned in adulthood. The study of phonetic drift, a phenomenon based in links between similar sounds, is also likely to improve our knowledge of cross-linguistic similarity and processes of diachronic sound change. In this way, the study of phonetic drift is significant not only for our understanding of language contact, but also for our understanding of language in general. Moreover, given how common the phenomenon will be demonstrated to be, it should be taken into consideration in the design and administration of language instruction and linguistic research.

1.2 Place in the Literature

The dynamics of interaction between two language systems have long been the subject of linguistic inquiry (see, e.g., Albert and Obler 1978; Grosjean 1982). Early observations of cross-language interaction between phonologies focused on the influence of the first language (L1) on the second language (L2) and the phenomenon of "foreign accent" as a result of two related assumptions: the existence of a so-called "critical period" for language acquisition and a unidirectional kind of cross-language influence. The classic view of the critical period (Penfield and Roberts 1959; Lenneberg 1967; Scovel 1969) holds that biological changes in brain development are responsible for the general decline in ability to learn another language with increasing age: children are better at acquiring language than adults because they have not yet passed this critical period of neural maturation. In contrast, adults become neurally rigid, such that the linguistic structures of L1 are eventually fossilized. It follows that while the L1 may cause some interference in adult acquisition of an L2, the L1 itself should not be affected.

More recent work in phonetics and second language acquisition has challenged both of these notions. First, some second-language researchers (e.g., Flege 1987a; Flege and MacKay 2010) have pointed out numerous problems with the basic enterprise of proving that a maturationally defined critical period exists (e.g., controlling for affective variables like motivation, language attitudes, and cultural affiliation) and have shown, moreover, that the critical period hypothesis is inconsistent with empirical findings such as a gradient decline in overall production accuracy with increasing age of L2 acquisition.[1] Second, a growing body of research (e.g., Pardo 2006; Nielsen 2007a; Babel 2009b; Kim 2009) has shown that within L1, phonetic representations can be quite malleable and quickly adjust to the phonetic characteristics of other L1 speakers under certain phonological and social conditions. Finally, there is mounting evidence that L1 can, in fact, be affected by the learning of an L2. Recognition of L2 influence on L1 goes back as early as Selishchev (1925), but is first discussed extensively in the work of Flege and colleagues (e.g., Flege 1987b, 1995, 2002; Flege, Schirru, and MacKay 2003; Flege 2007). In particular, Flege's Speech Learning Model proposes that "[p]honetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category" (Flege 1995:239). Thus, the notion of a static, fossilized L1 has largely been replaced with that of a dynamic and ever-changing L1.

Though it is clear that L1 speech accommodation to other L1 talkers occurs readily and rapidly, what remains unclear is how malleable L1 representations are in the face of L2 learning. These two situations—within-language convergence (or divergence) between L1 talkers on the one hand, and cross-language convergence (or divergence) between L1 and L2 sounds on the other hand—differ critically in that neither the sociolinguistic motivation to accommodate nor the lexical overlap supporting accommodation in the former case are present in the latter case. While there are clear social reasons why L1 talkers might come to speak more or less like other members of the same speech community (see, e.g., Giles, Coupland, and Coupland 1991; Chartrand and Bargh 1999), there is no similar motivation for L2 learners to modify their L1 representations with respect to L2: doing so changes nothing about the social distance between them and native speakers of the L2, given that L1 is not a shared language, nor could doing so accomplish a modification of the social distance between them and other L1 speakers in any intended way, given that the change is not motivated by L1 input (if anything, it might—unintentionally—increase social distance, as L2-influenced L1 speech is likely to be perceived by other L1 speakers as accented). Moreover, in the case of L1 speech accommodation there are several ways in which L1 tokens may be connected to each other that are not available in the case of L1-L2 phonological interaction in most models of bilingual speech processing

[1] However, note that other researchers, who have controlled more carefully for these affective variables, have found a clear relationship between acquisition outcomes and age of acquisition, which suggests that late acquisition of a language suffers at least in part because it occurs outside an early sensitive period for language acquisition. Even when late acquirers come to the table with virtually no prior linguistic experience that could possibly interfere in the acquisition process (deaf individuals not exposed to sign language in childhood), their acquisition outcomes still end up as worse than those of early acquirers and, for that matter, worse than those of late L2 acquirers (Mayberry and Eichen 1991; Mayberry 1993; Morford and Mayberry 2000; Mayberry, Lock, and Kazmi 2002; Mayberry and Lock 2003; Mayberry 2007, 2010).

and production (e.g., de Bot 1992; Paradis 2001). When one L1 English talker hears another L1 English talker utter the word *pot* [pʰɑt], for example, that token can link up to the first talker's previously experienced tokens of *pot* on multiple levels of linguistic representation—semantic, syntactic, and phonological. In contrast, when an L2 learner of Korean hears another Korean speaker utter the unfamiliar word 팥 [pʰat] 'red bean', that token can link up to previously experienced tokens of L1 words (e.g., *pot*) only on the phonological level.

Despite these differences between L1 speech accommodation and L2 learning, a number of controlled phonetic studies (e.g., Flege 1987b; Major 1992; Sancier and Fowler 1997) have shown that production of L1 categories shifts toward the phonetic norms of similar L2 categories when speakers have been living in an L2 environment for a long time. The framework of the Speech Learning Model (Flege 1995) analyzes this sort of change in L1 as arising from an "equivalence classification" of similar L1 and L2 sounds that results in their becoming perceptually linked, thereby allowing both sounds to be affected by input in L1 or L2. But this begs the question: when and how does this linkage of L1 and L2 categories begin to affect the production of L1 sounds? Previous studies on adult L2 learners have not been able to address this question because they investigate the pronunciation of L2 learners who are near the "end state" of L2 acquisition (i.e., highly proficient bilinguals who have spent years in an L2 environment). Because the literature skips over the period of adult L2 acquisition during which cross-language connections are likely established, it has not been able to address the question of whether cross-language category linkage is a phenomenon that occurs early or late in L2 development. Moreover, this work has focused largely on languages that are related and, moreover, share the same alphabet, representing many similar sounds with identical graphemes (e.g., English, French, Portuguese, Italian, Spanish). Thus, conclusions about phonetically based identifications of L2 sounds with L1 categories have been confounded by the orthographic relationship between the languages examined previously, especially in light of the prominent role that written representations typically play in formal L2 education. This confound makes it unclear whether cross-language identification of similar sounds is actually based upon the phonetic relationship between the sounds in contact or simply based on identity in orthographic representation. In short, while providing evidence of L1 phonetic drift during L2 development, the L2 speech literature has left two gaps: a temporal one (the early stages of L2 learning) and a structural one (unrelated pairs of L1 and L2).

Thus, while it is known that phonetic drift of L1 can happen, the nature and time course of this phenomenon are poorly understood. Previous studies, such as those conducted in the framework of the Speech Learning Model, claim that phonetic drift arises due to perceptual linkages between phonetically similar L1 and L2 sounds. This claim implies two things. First, it implies that speech production draws upon non-unitary sound categories encoding a distribution of variants, which may be shifted by overlapping variants from a different distribution associated with a close, but non-identical sound. Similar conceptions of phonetic representations recur throughout the speech production literature (as discussed in Chapter 2). Second, the idea of perceptual linkage between sounds implies that phonetic drift occurs at the level of segmental categories: L1 sounds are influenced by individual L2 sounds to which they have become perceptually linked. This view, however, is inconsistent with findings on phonetic drift in L1 vowel production, which have been shown to be difficult to explain in these terms (Guion 2003), suggesting that phonetic drift cannot be accounted for strictly in terms of links between pairs of L1 and L2 segments. With regard to the time course of phonetic drift, principles of the Speech Learning Model imply that phonetic drift may occur early in L2 acquisition. Equivalence classification is described as a mechanism of categorization, suggesting that it is an automatic process. If a novel sound encountered while learning an L2 is therefore classified as a similar L1 sound automatically, it follows that the potential for perceptual linkage between the L1 and L2 sounds and concomitant phonetic drift of the L1 sound will be virtually immediate. In short, based upon empirical findings of phonetic drift reported in the literature and theoretical statements about speech learning, it is reasonable to postulate that phonetic drift occurs via cross-language connections that extend beyond segmental parallelisms, and moreover, that it begins from the onset of L2 acquisition. If this were the case, it would mean that production of speech segments is a multifaceted task based on more than segment-sized representations, and that these representations—far from fossilizing—are constantly updated in adulthood.

1.3 Outline of the Dissertation

In this dissertation, the nature and time course of phonetic drift are reexamined with a focus on the very first weeks of native English speakers' immersion in a Korean language environment. In doing this, the current study broadens the scope of previous research on phonetic drift in two ways: (i) by investigating novice L2 learners, and (ii) by concentrating on a pair of languages that are unrelated both genetically and orthographically. Given that L1 phonological categories can be influenced by the phonetic characteristics of similar L2 categories, this study investigates whether they are affected from the very first stages of L2 acquisition or only at the later stages of L2 acquisition that have been previously examined. The null hypothesis in this case is that L1 categories remain unchanged in the short term, but this remains to be demonstrated empirically. Thus, the objective of the present study is to test the alternative hypothesis that, similar to the case of L1 speech accommodation, L1 phonetic drift during L2 learning occurs rapidly, from the very beginning stages of L2 learning. This hypothesis is tested in two case studies focusing on two different types of phonemic categories examined in the L1 accommodation literature: consonant categories differing in terms of articulatory timing (i.e., laryngeal categories) and vowel categories differing in terms of articulatory targets (i.e., vowel quality categories).

The dissertation is divided into seven chapters. Chapter 2 reviews several bodies of research supporting the basic premise of L1 speech malleability; models of L2 speech acquisition accounting for cross-linguistic phonetic phenomena; and empirical studies providing evidence for L2 influence on L1. It then motivates a set of specific predictions for a longitudinal study of novice L2 learners' L1 production. Chapter 3 describes the design of the longitudinal study, as well as the acoustic measurement protocols used in the analyses of learners' production over time. Chapter 4 presents the findings on phonetic drift in L1 consonants (laryngeal categories), while Chapter 5 presents the findings on phonetic drift in L1 vowels (vowel quality categories). Chapter 6 synthesizes the results of Chapters 4 and 5 with previous results found in the literature; discusses the findings with respect to the notion of cross-linguistic similarity, exemplar models of phonology, and historical change in contact situations; and considers the implications for language pedagogy as well as research standards in linguistic studies. Chapter 7 summarizes the main findings and suggests avenues for further research.

Chapter 2

General Background

2.1 Research Context

The extent to which human knowledge and behavior is determined by an innate genetic endowment ("nature") or developed via experience with the world ("nurture") has been the object of a longstanding debate in philosophy, biology, and the social sciences. In one version of a purely nativist view, biology is thought to be deterministic, such that individuals' behavioral characteristics are simply the reflex of their genetics (for a review, see de Melo-Martín 2004). In contrast, a purely empiricist view claims that humans start out as a blank slate, or *tabula rasa*, and that their behavior is the result of input and interaction with their specific environment (Locke 1690). Today, most scholars take an intermediate position, accepting that in a number of ways human beings are a product of both innate and experiential components, which furthermore interact with each other (e.g., Moore 2003; Ridley 2003).

In linguistics, the tension between "nature" and "nurture" has motivated a great deal of research into the questions of how humans come to master a language and how the process of language acquisition differs between childhood and adulthood. Work in first language (L1) acquisition has been concerned with understanding how a child acquires a complex linguistic system largely without explicit instruction—a question that has motivated research programs investigating the influence of an innately endowed Universal Grammar (Chomsky 1957; Bickerton 1984) as well as the role played by timely input (e.g., Curtiss 1977, but cf. Jones 1995) and distributional, probabilistic, and expectation-based analyses of the input in shaping L1 competence (Seidenberg and McClelland 1989; Saffran, Newport, and Aslin 1996; Saffran, Aslin, and Newport 1996; Hudson Kam and Newport 2005, 2009; Ramscar and Yarlett 2007).

Similarly, work in second language (L2) acquisition has examined the relative contributions of innate and experiential components in shaping L2 competence. One major body of research focuses specifically on the effect of preexisting linguistic knowledge on the acquisition of an L2 by adult learners. The ways in which the system established for one's L1 may influence the system established for an L2 have thus been of high interest to L2 researchers. Studies in L2 acquisition have repeatedly shown that the structure of an adult learner's L2 system, while approximating that of the target L2 to some degree, usually shows differences from the standard of native speakers that are attributable to influence from the learner's L1 system. This cross-language influence is often referred to as "interference" or "transfer" from L1 (see, e.g., Lado 1957), and in regard to L2 phonology, it forms the basis of having an identifiable foreign accent in L2.

The phenomenon of foreign accent—non-native-like pronunciation of L2 by speakers who typically acquired the L2 as adults—has been documented in a wide variety of L2 acquisition studies (for a broad overview, see Major 2001). Some of these studies focus on the effect of learner age on perceived foreign accent (e.g., Flege, Munro, and MacKay 1995b), while others are concerned with identifying which aspects of an L2 phonology are most likely to be produced as accented. Numerous studies have reported on learners' different degrees of success with various segmental categories and syllable structures (e.g., Yamada 1995; Broselow, Chen, and Wang 1998) as well as with suprasegmental properties of L2, such as lexical tone (e.g., Leather 1996), intonation (e.g., Tahta, Wood, and Loewenthal 1981), and rhythm (e.g., Kaltenbacher 1997). Two common conclusions of these studies are that "earlier is better" with respect to age of acquisition, and that aspects of the L2 phonology that are the hardest for learners to acquire in a native-like manner tend not to be those that are very different vis-à-vis L1, but rather those that bear close similarities to L1 (i.e., elements of the L2 phonology with L1 counterparts that stand to seriously "interfere" in L2). In short, prior L1 experience has consistently been found to play a role in L2 acquisition, a role that is mediated by both the amount of accrued L1 experience and the degree of cross-linguistic similarity between familiar L1 structures and unfamiliar L2 structures.

However, in addition to the effect of "nurture", something akin to "nature" has been found to play a role in L2 acquisition as well. While much of the L2 literature provides evidence of L1 transfer in L2 production, it has also been shown that an adult learner's L2 system tends to show traits attributable to neither L1 nor L2. These traits parallel features of L1 acquisition, and so are thought to be universal in nature, specified in whole or in part by Universal Grammar (White 1989). Historically this sort of data motivated a move away from simply analyzing learner language in terms of L1 transfer and instead toward looking at it as a semi-autonomous system called "interlanguage" (Selinker 1972). Interlanguage is thought of as the learner's dynamic system for L2 that incorporates L2 structure, L1 structure, and universally preferred structure in proportions that vary depending on a number of factors, including the time point in the acquisition process (e.g., the amount of L1 influence is generally found to decrease as learners become more proficient in L2; see Major 1987a).

Taken together, the findings of L2 acquisition studies suggest that, as in other domains of human behavior, both basic and experiential components play a role in L2 development. However, what is often assumed in discussions of L2 acquisition is that effects of experience are unidirectional, influencing L2 to the exclusion of L1. The assumption of L1 stability in adulthood is one that makes sense intuitively on analogy with biological traits such as height, which stops increasing after a certain point in an individual's life regardless of experience (e.g., amount and kind of food intake). Consequently, L2 studies that examine "forward interference" of L1 in L2 far outnumber those that discuss "reverse interference" of L2 in L1 (e.g., Felton 1990; Hussein 1994; Joseph 2009). Moreover, in the relatively few studies that specifically examine L2 effects on L1, there is the implication that these effects are detectable only after a large amount of L2 experience has accrued, as virtually all of this work focuses on proficient L2 learners. However, granting that one might expect the magnitude of L2 experiential effects to increase with amount of L2 exposure, it has never been empirically shown that there is in fact a threshold of L2 experience that must be passed in order for an L2 system to start affecting the L1 system. As such, it is conceivable that L2 actually begins affecting L1 early in L2 development, yet there are virtually no L2 acquisition studies that track learners from the onset of acquisition and, thus, no known data on this point. This gap in the literature is filled by the present study, which investigates the effects of elementary L2 experience on L1 production—in particular, the phenomenon of early PHONETIC DRIFT: subtle shifts in L1 sounds resulting from experience with similar, but non-identical L2 sounds.

By way of background, this chapter provides a broad overview of the literature on L1 speech malleability and L1-L2 interaction. In Chapter 1, it was argued that the assumption of a static L1 is invalid, and that L1 production is better characterized as malleable, the result of a dynamic linguistic system that continues to develop in adulthood. In this chapter, the assumption of a static L1 is first critically evaluated. Evidence is presented from studies of monolingual speech adaptation and speech accommodation demonstrating that L1 production is highly adaptive, sensitive to one's own feedback as well as to linguistic input from others. Subsequently, phonological accommodation of L1 to L2 is examined in a review of the literature on L2 speech learning and L1 attrition. An overview is provided of three influential models of L2 speech which summarizes the basic principles, predictions, and limitations of each model, with a special focus on the way in which these models view the relationship between dual phonological systems. A typology of dual phonological organization is then presented, providing the framework for a review of findings on L1-L2 phonological interaction in three groups varying in their level of exposure to and current use of L1 and L2: (i) L2-dominant speakers with extensive, but socially attenuated experience in an obsolescent L1 ("last speakers"), (ii) L2-dominant speakers with substantial, but interrupted exposure to L1 ("heritage speakers"), and (iii) typical "bilingual" speakers, ranging from L1-dominant late L2 learners to balanced bilinguals with early, extensive, and continuing experience in both L1 and L2.

The chapter concludes with a set of predictions regarding a longitudinal production experiment that constitute the point of departure for the case studies presented in Chapters 4–5. This experiment tests the hypothesis of temporal continuity in cross-language phonetic influence—that is, of L1 production changing not only in highly experienced, proficient L2 speakers, but also in novice, non-proficient L2 learners. As discussed in Section 2.5, it is predicted that phonetic drift of L1 production will occur early in the course of adults' L2 learning, and that this drift will result in the phonetic approximation of L2. Furthermore, L1 production changes are predicted to vary across different aspects of contrast in terms of generality, due to influence from L2 at different levels of L2 structure. Thus, as a study testing predictions about rapid L1 production adjustments to L2 input, the current study is informed by previous studies of rapid L1 production adjustments generally, as well as by studies of L1 production changes influenced by L2 experience specifically.

2.2 Malleability of L1 Production

Far from being static, L1 production has been shown to change in response to a number of external factors. This section summarizes relevant findings from three areas of research: change in the face of attenuated feedback, adaptation to altered feedback, and accommodation to the speech of native and non-native interlocutors. Despite the different goals and experimental paradigms of these sets of studies, what they have in common is the finding that L1 is malleable: production adjusts relatively quickly to input from the environment, rather than remaining the same.

2.2.1 Change with Attenuated Feedback

One of the earliest documented cases of externally influenced changes in L1 production is the so-called "Lombard effect" (Lombard 1911; Lane and Tranel 1971) whereby people raise their voices in noisy environments. Observing the parallelism between certain production changes in deaf people and similar production changes in hearing people faced with noise, Lombard found experimentally that subjects with fully intact hearing involuntarily raised their voices when the auditory feedback from their own speech was attenuated with masking noise, an effect that appears to arise not only from the need to make oneself audible to an interlocutor, but also from the need to simply hear oneself while speaking. The Lombard effect results in changes not only in amplitude, but also in fundamental frequency (f_0), duration, and spectral profile, including spectral tilt, energy balance, and formant frequencies (Summers, Pisoni, Bernacki, Pedlow, and Stokes 1988), although there is significant inter-speaker variability—especially between males and females—with respect to how the effect is realized (Junqua 1993). These acoustic shifts are accomplished in part by changes in speech breathing (Winkworth and Davis 1997) and, at higher noise levels, affect content words disproportionately relative to function words (Patel and Schell 2008).

The similarities between the Lombard effect on adults with normal hearing and the effects of hearing loss on adults deafened after learning to speak were explored in an extensive series of studies by Lane, Perkell, Svirsky, and colleagues. In one study by Lane and Webster (1991), it was found that, in comparison to the speech of hearing adults, the speech of post-linguistically deafened adults showed significant changes, many of which were reminiscent of the Lombard effect: deaf adults produced greater variability in pitch, higher mean pitch in both stressed and unstressed syllables, and overall slower speech, in addition to less differentiation of consonants in terms of place of articulation. Subsequent studies focused on post-linguistically deafened participants with cochlear implants, examining the effects of auditory feedback on speech production by comparing speech produced when the implant was either absent or turned off (depriving the speaker of auditory feedback) to speech produced when the cochlear device was implanted and turned on (providing the speaker with auditory feedback). Studies in this vein examined the effects of feedback on global properties of speech such as airflow, rate, sound level, and f_0 (Lane, Perkell, Svirsky, and Webster 1991; Perkell, Lane, Svirsky, and Webster 1992; Svirsky, Lane, Perkell, and Wozniak 1992; Lane, Perkell, Wozniak, Manzella, Guiod, Matthies, MacCollin, and Vick 1998) as well as on segmental categories such as vowels (Perkell et al. 1992; Svirsky et al. 1992; Lane, Matthies, Perkell, Vick, and Zandipour 2001; Vick, Lane, Perkell, Matthies, Gould, and Zandipour 2001; Perkell, Numa, Vick, Lane, Balkany, and Gould 2001), stops (Lane, Wozniak, and Perkell 1994; Lane, Wozniak, Matthies, Svirsky, and Perkell 1995), and sibilants (Matthies, Svirsky, Lane, and Perkell 1994; Matthies, Svirsky, Perkell, and Lane 1996).

With regard to global properties of speech, Lane et al. (1991) found that provision of auditory feedback resulted in significant changes in speech breathing. In a study examining three deaf speakers before and after they received cochlear implants, Lane et al. collected aerodynamic data via inductive plethysmography while the speakers read English passages. Following activation of their cochlear implants, all subjects were found to change their average airflow in approximation to normal levels: the two subjects with initially low airflow rates increased airflow, while the subject with an initially high airflow rate decreased airflow, with concomitant changes in air expenditure per syllable. In addition, one speaker went from ending breaths at air levels drawing upon "expiratory reserve volume" (i.e., breathing out too much) to ending them at closer to typical levels. These results were consistent with those of Perkell et al. (1992) and Lane et al. (1998), who used a similar longitudinal design.

In a study of four deaf speakers before and after they received cochlear implants (two of whom became deaf in adulthood and two of whom suffered severe hearing impairment in childhood), Perkell et al. (1992) also found marked changes in speech production due to activation of a cochlear implant. They measured changes in sound-pressure level (SPL), duration, amplitude, f_0, first formant frequency (F_1), second formant frequency (F_2), and spectral tilt (as measured by the difference between the first two harmonics of the spectrum, $H_1 - H_2$). In accordance with the findings of Lane and Webster (1991), it was found that there were "abnormalities of average parameter values and patterns of contrast" in the pre-implant speech of these deaf subjects (e.g., exaggerated differences in duration between tense and lax vowels, unusually high average f_0, exaggerated patterns of intrinsic f_0, reduced range of F_1 and F_2, unusually low $H_1 - H_2$). However, the provision of auditory feedback via the cochlear implant resulted in significant changes in subjects' speech, which tended to move in the direction of typical values: average SPL, duration, and f_0 all decreased for one or more subjects. Changes were also found in average F_1, $H_1 - H_2$, and airflow, although these were argued to follow from the observed changes in SPL, duration, and f_0. Notably, the subjects who were deafened in childhood patterned differently from the subjects who were deafened in adulthood. The latter group showed a stronger trend towards approximating typical values of speech parameters. This asymmetry led the authors to conclude that "production gains are governed at least as much by prior linguistic experience as by perceptual gains" from new auditory information.

The long-term effects of cochlear implant activation observed in Perkell et al. (1992) were compared to short-term effects in a study by Svirsky et al. (1992), who used a slightly different paradigm involving the deactivation and reactivation of the cochlear implant in three relatively experienced cochlear implant users (each of whom had been using their implant for at least six months). In this study, Svirsky et al. measured the same parameters analyzed in Perkell et al. (1992) in three conditions: (i) 24 hours after the subject's cochlear device had been deactivated, (ii) immediately after the device was reactivated, and (iii) immediately after the device was once again deactivated (following its 30-minute reactivation in the previous condition). Reactivation of the cochlear implant in the reactivation condition was found to have significant effects on subjects' speech, which were generally consistent with the long-term effects of initial cochlear implant activation in these subjects found by Perkell et al. (1992). Moreover, differences were observed across conditions in the general pattern of change, which for many parameters occurred more rapidly and to a greater degree in the reactivation condition than in the reactivation-deactivation condition. This disparity suggested that auditory feedback helped subjects make adjustments to articulatory parameters that had drifted in the absence of feedback, which then persevered when feedback was taken away again. Nonetheless, several parameters showed rapid changes in both the reactivation condition and the reactivation-deactivation condition. These findings thus indicated that auditory feedback aids not only in the long-term calibration of articulatory parameters documented in Perkell et al. (1992), but also in more short-term speech adjustments.

Consistent with these previous studies, Lane et al. (1998) found that the provision of auditory feedback resulted in significant changes in both speech breathing and sound level. In a study examining seven deaf speakers before and after they received cochlear implants, Lane et al. collected acoustic and aerodynamic data from the speakers' productions of the English vowels as well as a short passage. Their data showed that following the activation of their cochlear implants, nearly all speakers reduced their SPL. At the same time, several speakers increased glottal width (as measured by $H_1 - H_2$). Moreover, nearly all speakers went from ending their breaths at levels below "functional residual capacity" to ending their breaths at closer to typical levels, and their average expenditure of air per syllable approached typical levels as well. Thus, the results suggested that auditory feedback aids in economizing effort on the part of the speaker in two ways: it prevents speakers from talking unnecessarily loudly, and it also permits them to adjust their glottal configuration and respiration for maximally efficient airflow.

In addition to work on suprasegmental properties, several studies have examined the effect of auditory feedback on vowel articulation specifically. Temporal and spectral data in Perkell et al. (1992) and Svirsky et al. (1992) suggested that novel auditory feedback has a significant effect on vowel production, and this effect was further examined in a study by Lane et al. (2001), who compared vowel production and coarticulation in seven post-linguistically deafened adults and two normal hearing adults. Their results showed that the degree of anticipatory coarticulation (as measured by ratios of F_2 at vowel onset vs. vowel midpoint, ratios of F_2 in point vowels, and locus equations) remained largely the same in the deaf adults before and after activation of cochlear implants; moreover, the amount of coarticulation was similar to that found in the hearing subjects. In addition, there was no consistent change in vowel dispersion: as measured by vowel formants, four deaf subjects showed a decrease in dispersion, while the other three showed an increase. Nearly all, however, significantly reduced vowel durations.

In contrast, Vick et al. (2001) found for a group of eight post-linguistically deafened adults that activation of cochlear implants had a significant effect on both their perception and production of vowels, with corresponding changes in degree of intelligibility to hearing speakers. In this study, changes in perception and production were found to co-occur, such that speakers who produced less acoustic separation between neighboring vowels before activation often showed both an increase in their ability to perceptually distinguish neighboring vowels and an increase in the acoustic separation produced between vowels after activation. The enhancement of vowel contrasts in production subsequently resulted in improved intelligibility to listeners with normal hearing, suggesting that auditory feedback may play a role in articulatory adjustments made in the interest of accurate perception on the part of listeners.

The hypothesis that auditory feedback may influence production of vowel separation was further tested in a study of English- and Spanish-speaking cochlear implant users by Perkell et al. (2001). Perkell et al. hypothesized that although a speaker without auditory feedback might be expected to economize articulatory effort (resulting in reduced vowel separation), this effect should be manifested differently in languages with more or less crowded vowel systems (and, thus, more or less need to produce acoustic separation between neighboring vowels). Utilizing the reactivation-and-deactivation design of Svirsky et al. (1992), Perkell et al. found evidence that crowdedness of the vowel space does indeed mediate the effects of changes in hearing status on the production of acoustic vowel separation: while Spanish-speaking subjects showed variability in terms of changes in average vowel separation, English-speaking subjects were uniform in producing more vowel separation with auditory feedback than without. This finding was thus consistent with the results of Vick et al. (2001), who also found evidence of vowel contrast enhancement when auditory feedback was present.

Effects of auditory feedback on speech production have also been examined for stop consonants. In work by Lane et al. (1994), the production of English plosives was elicited from four post-linguistically deafened adults before and after they received cochlear implants, and these productions were analyzed for voice onset time (VOT) and syllable duration. As most speakers were found to produce shortened syllable durations after activation of their implants, measurements of VOT were adjusted with respect to differences from mean syllable duration, and these adjusted measurements for VOT were compared before and after implant activation. Results showed that before implant activation, the deaf speakers tended to produce VOTs that were too short compared to phonetic norms for hearing speakers; this was the case for all speakers' voiced plosives and two speakers' voiceless plosives. However, following implant activation, subjects were able to identify the voicing of plosives relatively accurately, and the majority also increased their adjusted VOT in approximation to normal values.

In a follow-up study by Lane et al. (1995), VOT and the "postural" parameters of SPL, f_0, and spectral slope were measured in five cochlear implant users' productions of English plosives, which were collected longitudinally before and after activation of their implants. Changes in VOT (adjusted for speech rate) found in this long-term experiment were furthermore compared to changes found in a short-term experiment using the same sort of design as Svirsky et al. (1992). In the long term, nearly all subjects were found to increase the adjusted VOT of voiceless and/or voiced plosives pre-activation to post-activation, with concurrent changes in SPL, f_0, and spectral slope. In the short term, similar results obtained for voiced plosives, but not for voiceless plosives, which did not change significantly in VOT with brief changes in hearing status. In this study, Lane et al. attempted to account systematically for the link between changes in speech posture and changes in VOT via regression analyses relating these two types of changes. The results of these regressions indicated that the two types of change were indeed related to each other, but primarily for voiced plosives. There were increases in VOT—both in the long term and in the short term, and especially in the case of voiceless plosives—that were not accounted for by changes in speech posture, suggesting that the presence of auditory feedback resulted in VOT adjustments separate from the adjustments in postural parameters.

Auditory feedback has also been shown to have a significant effect on production of sibilants. In Matthies et al. (1994), five cochlear implant users' production of English /s/ and /ʃ/ was examined at three time points (before implant activation, early after activation, and six months after activation), and pre- and post-activation speech samples of the subjects were examined in terms of how they were output by the cochlear device. Results showed that two subjects produced contrast between /s/

and /ʃ/ (in terms of spectral medians) well even pre-activation and continued to do so post-activation. The three other subjects showed reduced contrast pre-activation and increased the contrast post-activation with the aid of auditory feedback, which appeared to allow them to hear the differences in they were effecting in their production. In a subsequent study by Matthies et al. (1996), both acoustic data (spectral properties of the sibilants as well as general postural parameters) and articulatory data (from electromagnetic articulography) were collected for one cochlear implant user's production of sibilants while his implant was activated and while it was deactivated. Matthies et al. found that while the spectral profile of /ʃ/ productions was related to the position of the tongue blade, changes in spectral contrast between /s/ and /ʃ/ could not be accounted for by changes in general postural parameters. These findings were thus consistent with those of Matthies et al. (1994) in suggesting that auditory feedback aids in articulatory tuning of sibilant production.

It should be noted that improvements in production due to auditory feedback are substantial enough that they lead to general increases in intelligibility. This result was found by Vick et al. (2001), as well as by Gould, Lane, Vick, Perkell, Matthies, and Zandipour (2001), who examined changes in the speech intelligibility of eight post-linguistically deafened adults who received cochlear implants. These adults were recorded producing CVC words before activation of their implants and at six and twelve months after activation, and their productions were then played in noise to a group of seventeen listeners in a word identification task. Listeners' performance in this task pointed to a general trend of increases in intelligibility post-activation, although there was considerable inter-speaker variability, which was attributed to inter-speaker differences in a number of dimensions: level of hearing pre-activation, age of onset of hearing loss, and kind of hearing loss. Despite this variability, all eight speakers were found to improve in their vowel or consonant intelligibility, and five of the eight improved in both.

The manner in which auditory feedback is used in speech production is a question that runs through much of the work discussed above, as well as the work on speech adaptation discussed below. As summarized by Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico, Wozniak, and Guiod (1997), given that auditory feedback cannot be transmitted and processed quickly enough to be used in real-time adjustment of the individual gestures of segmental articulations, it seems that auditory feedback must be involved in tuning an "'internal model' of the relation between articulatory commands and acoustic results". This implies, then, that auditory feedback is used in the process of speech production for two purposes: "to help acquire and then maintain parameter settings of the internal model for segmental control, and to provide information for the regulation of some suprasegmental factors that influence intelligibility", such as SPL, f_0, and rate (which, as discussed above, were found to be adjusted very rapidly).

In sum, studies involving attenuated auditory feedback indicate that L1 speech production is sensitive to the presence of auditory feedback, which seems to play an important role in both short-term and long-term calibration of speech production.

The results of studies on the Lombard effect show that talkers with normal hearing amplify their production in compensation for decreases in auditory feedback from their own speech. Moreover, the results of studies on auditory deprivation in post-linguistically deafened adults suggest that absence of auditory feedback leads to a drift away from L1 speech norms, while provision of auditory feedback allows deafened adults to recalibrate their production in approximation to typical values. Besides providing evidence that L1 speech production is malleable, the work described above is relevant to the present study because it suggests that that there are fundamental differences between speech parameters related to segmental control (e.g., VOT) and speech parameters related to more global, "postural" aspects of production (e.g., f_0, rate, SPL). Such a dichotomy might thus be expected to exist also in the phonetic drift of L1 speech parameters under influence from the different settings of those parameters in L2.

2.2.2 Adaptation to Altered Feedback

Effects of taking away auditory feedback are similar in many ways to effects of delaying auditory feedback. Situations in which auditory feedback is delayed occur in everyday life when the acoustic signal is amplified (drowning out the real-time, non-amplified feedback from one's voice), but returned with a delay (e.g., when one talks into a loudspeaker system or makes an international phone call). Lee (1950) observed that delayed auditory feedback results in several kinds of changes to production. He noted that when faced with this sort of non-canonical feedback, some speakers "develop a quavering slow speech of the type associated with cerebral palsy; others may halt, repeat syllables, raise their voice in pitch or volume"; furthermore, the situation can be quite frustrating and/or tiring for speakers, who sometimes "reveal tension by reddening of the face". Different effects were observed for different speech genres as well as for production in L1 versus L2, though in no case were speakers able to completely ignore the feedback delay. Yates (1963) enumerated a number of specific effects on production—"prolongation of vowels, repetition of consonants, increased intensity of utterance"—that are consistent with the descriptions of Lee (1950). In addition, there were significant individual differences in how speakers coped with feedback delay: some were found to "show little disturbance", while others were "almost totally incapacitated". These individual differences in effects of delayed feedback appear to be related to certain personality variables. Citing the work of Spilka (1954), Yates observed that increases in vocal intensity variance in response to delayed feedback were correlated with "strong negative self-attitudes, poor personality adjustment, and paranoid tendencies"; on the other hand, decreases in vocal intensity variance were correlated with "schizoid modes of behavior".

Though delayed feedback is clearly disruptive to speech production, it is essentially veridical; it is the timing of the acoustic signal's reception by the speaker, rather than the information contained in the signal itself, that is altered, affecting the speaker's production. The effects of altering the information contained in auditory feedback have also been explored in research on L1 adaptation. Like research on cochlear implant users, research on adaptation in speech has sought to understand how people make use of auditory feedback in modulating their speech production. However, rather than manipulating the quantity of feedback, adaptation studies have generally tried manipulating the quality of feedback that a talker hears (i.e., transforming the information contained therein) and observing the changes that subsequently occur in the talker's speech. Work in this paradigm has examined the effects of altered auditory feedback on the production of two main acoustic properties: f_0 (Elman 1981; Kawahara 1993, 1994; Burnett, Senner, and Larson 1997; Burnett, Freedland, Larson, and Hain 1998; Larson, Burnett, Kiran, and Hain 2000; Jones and Munhall 2000, 2005) and vowel formants (Houde and Jordan 1998, 2002; Purcell and Munhall 2006; Pile, Dajani, Purcell, and Munhall 2007; Katseff and Houde 2008). The basic methodology used in most of these studies involves taking the acoustic signal of a talker's speech, transforming it in one or more ways in a signal processor, and returning it to the talker through headphones with a delay short enough that it goes unnoticed (and at a volume loud enough that the real-time, unaltered auditory feedback from the talker's voice is masked). In this way, talkers can be "tricked" into being influenced by distorted feedback that does not accurately reflect what they actually produced.

In the first work on effects of altering f_0 in auditory feedback, Elman (1981) found that distortions to feedback f_0 resulted in compensatory changes in five English speakers' produced f_0. In Experiment 1, subjects listened to a steady-state, target [ɑ] at one of eight different f_0 levels and then attempted to shadow the target while feedback from the subject's production was played over headphones (either unaltered, shifted up in f_0, or shifted down in f_0). Results of this experiment showed that while subjects had no trouble maintaining a constant f_0 with unaltered feedback, they tended to shift their f_0 in opposition to the direction of f_0 shift in altered feedback such that the feedback they heard was consistent with the target. Thus, when feedback f_0 was shifted up, subjects' produced f_0 went down, and vice versa, even despite subjects often being unaware of the feedback alterations. In Experiment 2, subjects completed a more linguistic task in which they repeated the target phrase *Where were you a year ago?* ten times with exaggerated intonation and as much consistency as possible. The task was completed twice while subjects received either unaltered or altered feedback. Similar to the results of the [ɑ] production experiment, results of the sentence production experiment showed more variability in f_0 across productions with altered feedback than with unaltered feedback, as well as some lengthening.

These patterns of compensation for f_0 shifts in feedback were also found in the work of Kawahara (1993, 1994) on Japanese. In an experiment similar to that of Elman (1981), subjects attempted to produce the Japanese vowel /a/ with a constant pitch, except on their own instead of in response to a stimulus. They completed this task while receiving unaltered or altered feedback. Results showed that subjects changed their f_0 when hearing altered feedback, and that the f_0 fluctuations were due to feedback shifts rather than to natural f_0 variation (Coleman and Markham 1991). Moreover, the corrective response to f_0 alteration occurred relatively quickly, with latencies of 100–200 ms.

A series of studies by Burnett, Larson, and colleagues further investigated the effects of altered f_0 feedback using a much larger group of subjects. Burnett et al. (1997) replicated the above results using a design similar to Kawahara (1994) and a group of 67 subjects. In this study, subjects produced steady-state vowels (/ɑ/) at their habitual speaking pitch and sang musical scales while attempting to ignore shifts in feedback f_0. These subjects were overwhelmingly found to be affected by f_0 alterations, with 96% increasing their produced f_0 in response to a decrease in feedback f_0 and 78% decreasing their produced f_0 in response to an increase in feedback f_0. In addition, response latencies were similar to those found by Kawahara (1993), ranging from 104 to 223 ms. In follow-up experiments by Burnett et al. (1998), it was found that the f_0 adaptation response was not significantly affected (in either magnitude or latency) by the overall intensity of the feedback, the presence of pink masking noise, or the magnitude of the f_0 shift. On the other hand, increases in the magnitude of the f_0 shift were correlated with a decrease in the proportion of adaptation responses opposing the shift and a corresponding increase in the proportion of responses following the shift, suggesting that there are two types of adaptation modes that may be used depending upon whether comparison is made to an internal or external f_0 reference and whether the disparity with the reference f_0 is large or small. Furthermore, longer durations of f_0 shift beyond 100 ms were associated with adaptation responses of longer duration and greater magnitude, as well as with two-peaked responses, suggesting that there may be two adaptation responses with different latencies. Larson et al. (2000) later found that the adaptation response was also sensitive to the velocity to the peak of the f_0 shift: as velocity of the shift increased, so did velocity of the response, while the magnitude of the response was found to decrease.

In addition to direct adaptation to f_0-altered feedback, Jones and Munhall (2000) found evidence of lasting sensorimotor adaptation of the sort found in Houde and Jordan (1998, 2002) for formant-altered feedback. In this study, eighteen subjects produced vowels in three conditions varying in f_0 alteration: unaltered feedback, upshifted feedback, and downshifted feedback. In the altered feedback conditions, feedback was incrementally shifted without the subjects' awareness; afterwards, it returned to normal. As in Elman (1981), subjects modified their f_0 in compensation for f_0 shifts in altered feedback. However, subjects continued to show this compensatory pattern after feedback returned to normal: when feedback went from upshifted back to normal, subjects increased their f_0, whereas when feedback went from downshifted back to normal, they decreased their f_0. These findings suggested that subjects modulated their f_0 production using not only the proximal auditory feedback directly, but also a (modified) internal model of pitch mapping motor commands to auditory feedback. When auditory feedback was altered, subjects quickly re-learned these mappings, associating the motor commands for a given f_0 with the acoustic feedback of a different f_0 (a higher f_0 when feedback was shifted up in f_0, for example). Consequently, when auditory feedback returned to normal following the upshifted condition, subjects—still using their new mappings—perceived the f_0 of their now-veridical acoustic output as lower than intended, causing them to increase their f_0.

Jones and Munhall (2005) followed up on their original study with an extension to lexically meaningful f_0—namely, Mandarin tone production. In the same sort of experiment, Mandarin-speaking subjects produced two Mandarin items with different tones: /ma˥/ 'mother' with the high level tone (Tone 1) and /ma˧˥/ 'hemp' with the mid rising tone (Tone 2). The f_0 of auditory feedback was gradually increased until the altered f_0 was one semitone above the real f_0. In response to this increase in heard f_0, subjects compensated by lowering their produced f_0. As in Jones and Munhall (2000), the adaptation in production persevered when feedback returned to normal, at which point speakers increased their produced f_0. Furthermore, adaptation was found to generalize to the production of a tone category different from the one subjects originally produced during feedback alteration, although adaptation was more robust for the original tone. These findings were consistent with those of Jones and Munhall (2000) in showing aftereffects of brief exposure to altered auditory feedback, suggesting a rapid remapping of relationships between motor commands and auditory feedback that was mediated by articulatory target.

Sensorimotor adaptation to altered feedback has also been examined with respect to the production of vowel quality. Houde and Jordan (1998, 2002) investigated the effects on vowel production of altering the formant structure of auditory feedback. To minimize feedback from bone conduction, they focused on whispered speech elicited in a two-hour production experiment. In the experiment, subjects (eight male American English speakers) produced CVC words containing the vowel /ɛ/ while wearing headphones—first in a baseline phase with no feedback alteration, then in a ramp phase in which feedback was altered gradually up to a target level of perturbation, next in a training phase in which feedback remained altered at the target level of perturbation, and finally in a test phase in which feedback returned to normal. Feedback was altered by sending the acoustic signal through a digital signal processor, which analyzed the formant structure to synthesize a new signal with formants (primarily F_1 and F_2) shifted either in the direction of [i] or in the direction of [ɑ]. The altered feedback was then fed into the subject's headphones with a 16-ms delay, which went unnoticed by subjects. There were three sets of stimuli: a training set consisting of items with the vowel /ɛ/ in a bilabial context (*pep*, *peb*, *bep*, *beb*), a test set consisting of items with the vowel /ɛ/ in non-bilabial contexts (*peg*, *gep*, *teg*), and a test set consisting of items with other vowels (*pip*, *peep*, *pap*, *pop*). Subjects received altered feedback only when they produced training items in the ramp and training phases; otherwise, they heard masking noise after the baseline phase. Subjects were thus blocked from ever hearing intermittent accurate feedback that could be used to recalibrate drifting production. In the end, subjects were generally found to change their vowel production in compensation for the formant shifts in altered feedback. In addition, the pattern of adapted production persevered with the return to unaltered feedback and generalized both to the other phonetic contexts for the same /ɛ/ vowel and to the other vowels in the test set items. These results were consistent with the findings of Jones and Munhall (2000, 2005) for f_0. However, considerable inter-speaker variability was found, with some subjects showing little adaptation and others showing nearly complete adaptation.

Purcell and Munhall (2006) combined elements of Houde and Jordan's (1998) design and Kawahara's (1993) design to conduct a study of real-time adaptation to formant-altered feedback. Whereas Houde and Jordan (1998) focused on the net effects of gradual formant alteration to a different vowel percept, as it was not their objective to test how quickly compensation for this alteration occurred, the study of Purcell and Munhall (2006) differed by shifting formants suddenly. Specifically, subjects pronounced steady-state vowels (as in f_0-alteration studies), to which alterations of F_1 were sometimes applied. The vowels pronounced were /ɪ, ɛ, æ/. On certain /ɛ/ trials, F_1 in auditory feedback was shifted up or down to result in the percept of either /æ/ or /ɪ/. Subjects were found to respond to these F_1 alterations by changing their production in a compensatory manner, although the magnitude of average F_1 compensation was relatively small (no greater than 11% and 16% of the F_1 shift for /ɪ/ and /æ/, respectively). Similar to the adaptive response to f_0 alteration, the adaptive response to F_1 alteration occurred relatively quickly (sooner than 460 ms after the onset of the alteration), suggesting that auditory feedback is used similarly in the control of f_0 and F_1.

Generalization of this adaptation was examined by Pile et al. (2007) in a study similar to that of Houde and Jordan (1998), except that subjects phonated normally instead of whispering. As in previous studies of formant-shifted feedback, subjects in this study produced /ɛ/, which was altered in both F_1 and F_2 to sound like /æ/. In response to this altered feedback, subjects again changed their production in a compensatory manner. In addition, the adaptation was again found to persist when feedback returned to normal, at which point subjects increased their produced F_1. However, in the process of "unlearning" associations with altered feedback, subjects actually increased their F_1 beyond their baseline F_1, although by the end of the experiment it was not significantly different from baseline. In contrast to the generalization found in Houde and Jordan (1998), subjects in this study did not show generalization of adaptation for /ɛ/ to other vowels (/ɪ/ and /e/). Moreover, associations with altered formant feedback were eventually unlearned regardless of what was produced right after feedback returned to normal: subjects who produced /ɪ/, subjects who produced /e/, and subjects who simply waited in silence all produced F_1 values that were not different from baseline by the end of the experiment. In short, although adaptation to formant-shifted feedback persisted for some time after the feedback shift was removed, it appeared to have only temporary effects on subjects' speech production.

Observing that compensation for altered feedback is typically incomplete, Katseff and Houde (2008) investigated the degree of compensation for F_1-shifted feedback in three experiments varying the size of F_1 shift. In Experiment 1, which shifted F_1 by as much as 250 Hz, it was found that for most of the seven subjects the amount of compensation followed the amount of feedback shift linearly and then began to level off at a shift of approximately 150 Hz. Experiment 2, which was the same length but decreased the maximum feedback shift to 90 Hz, confirmed that the leveling off of compensation found at the end of the first experiment was not due to time, but rather to the magnitude of feedback shift. Finally, results of Experiment 3, which examined the phonetic space of subjects' vowel categories with unaltered feedback, showed that degree of compensation decreased when compensatory productions began to leave the normal phonetic space of that vowel category. Taken together, the findings suggested that degree of compensation for altered feedback becomes less complete with larger feedback shifts, and that this may be due to the relatively large discrepancy between auditory and somatosensory feedback that obtains when productions migrate outside of the normal category space.

Evidence of the grounding effect of somatosensory feedback in speech production was documented in work by Tremblay, Shiller, and Ostry (2003) and Larson, Altman, Liu, and Hain (2008). Tremblay et al. (2003) attempted to dissociate the influence of somatosensory feedback from that of auditory feedback (which usually changed in conjunction with somatosensory feedback in previous studies) by altering somatosensory feedback alone. In this study, subjects producing the utterance [siæt] received altered somatosensory feedback via the application of a mechanical load to the jaw that changed its movement path depending on the velocity of jaw movement. Their production was found to adapt to the presence of the load, both in a vocalized speech condition (in which the presence of the load was found to have no systematic effect on the auditory feedback) and in a silent speech condition (in which auditory feedback was absent). Thus, these findings suggested that somatosensory feedback has its own effect on speech production and, moreover, that speakers have somatosensory goals independent of acoustic goals. While Tremblay et al. (2003) altered somatosensory feedback, Larson et al. (2008) removed it via anesthesia. In a study of adaptation to f_0-shifted feedback, they anesthetized the vocal folds of seventeen subjects so that somatosensory feedback would not be available to them while they completed the sort of production task used in previous studies of f_0-altered feedback. In comparison to a control condition in which the vocal folds were not anesthetized and somatosensory feedback was available, production in the anesthetized condition showed larger adaptive responses to altered feedback, suggesting that somatosensory feedback plays an important role in modulating the degree of adaptation to altered feedback.

In short, studies of adaptation to spectrally altered auditory feedback show consistently that L1 speech production compensates for feedback distortions in a rapid manner. The compensation modifies an internal model of relations between motor commands and auditory feedback, which takes some time to retune to normal feedback. However, compensation for auditory feedback alterations is typically incomplete, never going beyond 50% of the feedback shift—a finding that appears to be due to the simultaneous effect of somatosensory feedback. These findings are relevant to the study of L2 influence on L1 because they suggest that L1 speech production is adjusted whenever there is a mismatch between auditory feedback and a talker's internal targets for production. A mismatch between L2 auditory input and a talker's internal targets for L1 production is the norm in the case of L2 learning. It follows that if L2 auditory input is experienced in an analogous loop as L1 auditory input, talkers might not be able to help adjusting their L1 targets to L2 input, since this input, as human speech, is in a way still relevant to their L1 speech production, even if it occurs in a different language. Consequently, L2-influenced adjustments in L1 production might be expected to occur rapidly, yet incompletely, and to show both generalization and specificity, in the same manner that adaptation occurs to altered auditory feedback.

2.2.3 Accommodation to Another Talker

That talkers adapt not only to feedback from their own speech, but also to the speech of others has been amply shown in studies on both native and non-native talkers. Research on L1 speech accommodation has approached the question of how talkers come to sound like each other in two main ways: perception tests of the degree to which any modifications of talkers' speech over the course of a communicative interaction are perceptible to listeners, and acoustic analyses of talkers' productions before and after exposure to the speech of another talker. With regard to acoustic studies, while work in the area of speech adaptation to altered feedback has focused almost exclusively on frequency components such as f_0 and F_1, the speech accommodation literature has examined both frequency components (e.g., Pardo 2009) and aspects of articulatory timing (e.g., Nielsen 2008).

One type of accommodation to non-native talkers occurs in a variety of speech referred to as "foreigner talk" or "teacher talk" depending on whether the conversational context is instructional or not (Ferguson 1975, 1981; Freed 1981; Hatch 1983; Chaudron 1983; Long 1983). Foreigner talk may be thought of as "a register of simplified speech...used by speakers of a language to outsiders who are felt to have very limited command of the language or no knowledge of it at all" (Ferguson 1975:119), though not all modifications made in foreigner talk may necessarily be characterized as simplifications. Differences between speech directed to non-native speakers and speech directed to native speakers have been observed on all linguistic levels, including discourse structure, lexis, syntax, and morphology. For example, foreigner-directed speech tends to be characterized by more repetition and restatement, more topicalization, increased use of tag questions, decreased use of pronouns and ellipsis, decreased syntactic complexity, shorter mean utterance length, use of high-frequency words, and avoidance of idioms (Hatch 1983; Gass 1997). Phonological features of foreigner talk include slower speech rate, increased amplitude, less articulatory reduction, more frequent and longer pauses, and increased use of emphatic stress. These modifications vis-à-vis native-directed speech are thought to be made for the sake of facilitating communication with non-native speakers of lower proficiency in the shared language, and variation in proficiency among non-native speakers is reflected in significant variation in foreigner talk by native speakers (Gass and Varonis 1985; Roche 1998).

Speech accommodation also occurs in native-native interactions. On the phonetic level, accommodation to native talkers was first documented in studies of overall accent and paralinguistic properties of speech (e.g., amplitude, f_0), which found that talkers tended to converge in these dimensions over the course of a conversational interaction (Giles 1973; Natale 1975; Gregory and Hoyt 1982; Gregory 1990). Later work found instances of both convergence and divergence in overall accent as well as individual acoustic properties such as VOT and vowel formants (e.g., Bourhis and Giles 1977; Nielsen 2008; Pardo 2009). These accommodative adjustments to a talker's production have been accounted for either in terms of communication accommodation theory (Giles et al. 1991; Shepard, Giles, and LePoire 2001), in which different types of accommodation are predicted to arise from different conversational roles and social motivations in a communicative interaction, or in terms of an interactive alignment model (Pickering and Garrod 2004a,b), in which convergence between talkers follows from automatic priming activating structures in one talker that are used by the other talker.

Some of the earliest observations of speech accommodation within L1 come from the work of Giles and colleagues. Giles (1973) examined accent shift by conducting sociolinguistic interviews about the same topic (attitudes regarding crime and capital punishment) that paired thirteen working-class teenage interviewees from Bristol, England with each of two different interviewers: an adult speaker (Giles) with standard Received Pronunciation (RP) and a teenage speaker with a Bristol accent. Pairs of interview excerpts spoken by the same interviewee, one from each interview, were then played to a group of Bristol listeners and a group of Welsh listeners for evaluation. Listeners were asked to identify which sample contained a "broader" accent and less formal lexical/grammatical usage; they were also asked to rate on an eight-point scale the extent of accent change and grammatical change between the two samples. The results of this rating task showed that listeners were able to perceive changes in both accent and usage between the two interview situations (although Welsh listeners were generally more accurate in their accent evaluations than Bristol listeners). Specifically, interviewees tended to use less broad pronunciation and more formal usage with the RP interviewer than with the Bristol interviewer, a result that was interpreted as interviewees' standardizing their speech patterns for the sake of social approval from a higher-status interlocutor.

While Giles (1973) found evidence of "upward convergence" to a high-status talker, Bourhis and Giles (1977) found evidence of both convergence and divergence in an investigation of accent mobility in British English-speaking learners of Welsh. Like the study of Giles (1973), this study conducted interviews and then presented recordings of interviewees' speech to two untrained listeners for evaluation on an eleven-point accent rating scale. However, unlike Giles (1973), who investigated the response of one group of interviewees to two different interviewers, Bourhis and Giles (1977) examined the responses of two groups of interviewees to the same interviewer. In this study, an English interviewer talked to two groups of Welsh learners, who differed with respect to their motivations for learning Welsh: the first group was learning it largely for personal reasons, while the second group was learning it mostly for professional reasons. During the course of interviewing both groups, the English interviewer questioned the contemporary utility of Welsh, and the groups responded to this somewhat threatening commentary on the Welsh language in different ways. While the group with pragmatic motivations for learning Welsh was found to converge with the interviewer, the group with personal motivations for learning Welsh (i.e., the group with the stronger Welsh identity) was found to diverge from the interviewer by producing Welsh-accented English, suggesting that patterns of phonetic accommodation may serve as a means of emphasizing or de-emphasizing social differences between talkers.

Accommodation occurs even without social interaction, as demonstrated in a shadowing study by Goldinger (1998), who tested the effects on accommodation of word frequency, shadowing delay, and talker variance in three sets of production and perception experiments. Similar to Giles (1973), Goldinger (1998) measured the extent of accommodation produced by talkers via perceptual judgments from a separate group of listeners, except using an AXB similarity judgment task instead of an accent rating task. In the first production experiment, talkers produced a set of words varying in frequency three times: once during a shadowing task in which they shadowed the productions of a model talker (played over headphones), and once before and after the shadowing task (by reading the words aloud with no auditory input). In the second production experiment, talkers shadowed a set of non-words that varied in frequency of presentation during an initial training phase in which talkers were familiarized with the items. In the third production experiment, talkers also shadowed non-words, but with variation in the model talker they were shadowing. All three production experiments contained two conditions differing in shadowing delay—an immediate-shadowing condition with no delay between model production and shadowed production, and a delayed-shadowing condition with a 3–4 second delay between the two. The results of these three experiments suggested that frequency, shadowing delay, and talker variance all affected degree of phonetic convergence. Specifically, more convergence was found for lower-frequency words and less frequently presented non-words; with no shadowing delay; and with a constant model talker. These findings were consistent with an episodic theory of lexical access in which lexical representations comprise detailed memory traces of one's experiences with a particular word (rather than a unitary, abstract construct).

Elements of Giles's (1973) and Goldinger's (1998) studies were recombined in several later studies of phonetic accommodation. In a study of phonetic accommodation in American English speakers, Pardo (2006) combined an interactive, conversational task reminiscent of the interview situation set up by Giles (1973) with the AXB similarity judgment task used by Goldinger (1998). In the conversational task, one talker (the "giver"), who had a map with a path drawn on it through various landmarks, conversed with another talker (the "receiver"), who had an identical map with only landmarks, in an effort to direct the receiver along the given path. Pre-task and post-task productions of lexical items related to the task were recorded by both talkers and later played, along with repetitions of these same items that naturally occurred during the task, to a new group of participants in a perception experiment testing for phonetic convergence between talkers engaged in the task. Listeners in the perception experiment completed an AXB similarity judgment task in which they indicated whether X (production of one talker) was more similar to A or B (productions of the talker's partner at various points in the production experiment). Results of this experiment suggested that overall there was significant phonetic convergence between talkers in the task (both in same-sex pairs as well as mixed-sex pairs, discussed in Pardo 2009), but that the extent of convergence was affected by both talker sex and conversational role. Male pairs were judged to have converged more often than female pairs; moreover, while female talkers converged to receivers, male talkers converged to givers. These findings provided evidence that social aspects of a conversational situation mediate between perception and production and suggested that convergence arises not from automatic priming, but from effects of "entrainment" in a coordinated dynamical system in which two talkers may be paired with varying degrees of coupling and dominance.

A follow-up study by Pardo (2009) provided acoustic analyses of participants' speech in Pardo's (2006) study. Acoustic measurements of f_0 and duration were taken on the stimuli used in the perception experiment and translated into difference measures (specifically, differences between members of a pair of stimuli), and these difference measures were used as predictor variables in linear-regression modeling of listeners' similarity judgments. The predictor variables were found to account for 41% of the variance in judgments on female talkers and 7% of the variance in judgments on male talkers, suggesting that f_0 and duration played a larger role in convergence between female talkers than convergence between male talkers. Pardo also reported on acoustic data (F_1 and F_2) from pre- and post-task productions of the vowels /i, u, æ, ɑ/ in words unrelated to the map task in Pardo (2006). Comparisons of pre- and post-task productions showed either convergence or divergence depending on the type of vowel and the conversational role of the talker in the map task: high vowels converged, while low vowels diverged, attributable to more centralization (in particular, of the low vowels) for givers than for receivers. Crucially, these vowel data came from novel words that were not used during the task, suggesting that convergence to an interlocutor was generalized at a phonological level rather than confined specifically to words shared between talkers in the map task.

Generalization of accommodation also occurs in the domain of VOT, as found in work by Nielsen (2005, 2006, 2007a,b, 2008) following previous findings of VOT imitation (Fowler, Brown, Sabadini, and Weihing 2003; Shockley, Sabadini, and Fowler 2004). In a study of response times in a shadowing task, Fowler et al. (2003) conducted shadowing experiments contrasting "simple" and "choice" responses to VCV stimuli. In the simple task, subjects had to respond to the CV portion of a stimulus (one of /pa, ta, ka/) with one unchanging CV response, while in the more difficult choice task, subjects had to respond to the CV by shadowing that particular CV. Comparisons of response times in these two tasks showed that subjects responded to the CV in the choice task only slightly more slowly than in the simple task, which was taken as evidence that they were able to extract information from the signal for the purposes of articulation without an intermediate process of decision-making. Importantly, in one of the experiments VOT in the CV portion of the stimulus was lengthened, which resulted in subjects responding with lengthened VOT, although the lengthened response VOT did not precisely follow the lengthened stimulus VOT. This latter result was consistent with the finding of Mitterer and Ernestus (2008) showing that while Dutch speakers reliably shadowed the phonologically relevant distinction of presence vs. absence of stop voicing, they did not imitate phonologically irrelevant distinctions in duration of voicing during closure. Like Fowler et al. (2003), Shockley et al. (2004) also found that talkers in a shadowing task imitated lengthened VOT to a significant degree. Moreover, in an AXB similarity judgment task, talkers' imitations of lengthened-VOT stimuli were judged to be more similar to the stimuli than were baseline productions, with no influence of AXB presentation order.

In similar work, Nielsen (2005, 2006, 2007a,b, 2008) further examined VOT imitation, but specifically included novel stimuli to test for generalization. As in Shockley et al. (2004), subjects were exposed to model speech (and recorded pre- and post-exposure productions of a word list); however, in this case they simply listened to the model speech and were not asked to shadow it. Nielsen (2005) tested how subjects would produce words after exposure to stimuli with lengthened VOT to see if VOT imitation would generalize to new items. During exposure to the model speech, eight English-speaking subjects only heard items with initial /p/ that had lengthened VOT (in addition to fillers), while in the elicited production task, they read words that included items heard during exposure, new items with initial /p/, and new items with initial /k/, which they did not hear at all during exposure. Results showed that as in Shockley et al. (2004), subjects imitated the lengthened VOTs to a significant degree. However, not only did they produce lengthened VOT in /p/-initial words they were exposed to, they also produced lengthened VOT in new /p/-initial words as well as in /k/, indicating generalization at a subphonemic level (perhaps targeting the feature [+spread glottis]). Moreover, lexical frequency of the target words was not found to have an effect on the VOT imitation. Nielsen (2006) extended this study to seventeen subjects and found the same result, again with no effect of lexical frequency. Furthermore, in this case there was an advantage of previous exposure, with significantly more imitation in items that subjects were exposed to in model speech than in novel items. However, this advantage was not found in the later work of Nielsen (2007a,b, 2008), which took this line of research one step further by investigating imitation of shortened and lengthened VOTs, adding the type of VOT manipulation subjects were exposed to as a between-subjects variable. In this later work, a marked asymmetry was found between the two types of VOT manipulation: while there was significant imitation in the lengthened-VOT condition, there was none in the shortened-VOT condition (even though the shortened VOTs more closely resembled VOTs actually produced by subjects before exposure), a result attributed to the phonological pressure to maintain contrast with the voiced stops, which are characterized by short-lag VOT. As mentioned above, neither lexical frequency nor word-level overlap were found to have significant effects on imitation. Thus, the findings suggested that phonetic imitation occurs at a subphonemic level, but is nevertheless constrained by phonemic considerations.

Like Nielsen, Babel (2009b) used the shadowing paradigm of Goldinger (1998), complete with actual shadowing, in an acoustic study of vowel imitation. She was concerned both with the automaticity of spontaneous imitation and its mediation by social factors. Subjects in this study, White American English speakers, produced words before and after they shadowed the speech of a model talker, either a White American male or a Black American male. In some conditions, subjects were also exposed to a picture of the model talker, while in other conditions, subjects were only exposed to his speech. In addition, subjects completed an implicit association task measuring their racial bias in which they made speeded judgments on word stimuli while the four words *black*, *white*, *good*, and *bad* were displayed in various configurations within their field of vision. Results showed evidence of accommodation to the model talker in the task, but not across the board. Significant changes in production were localized mainly to the F_1 of the low vowels /æ, ɑ/, a result attributed to the broader range of low-vowel exemplars that can be drawn upon in imitation due to the particularly pronounced effects of prosodic structure on the realization of these vowels in terms of jaw height. Moreover, racial biases found in the implicit association task, as well as judgments of the attractiveness of the model talker, were found to have a slight, but significant, effect on the amount of imitation found in the shadowing task. Overall, the findings were consistent with the interpretation that subjects limited themselves to tokens within their phonetic repertoire in adjusting their productions towards the model talker's, and that these adjustments were mediated by "implicit socio-cognitive biases".

The focus of phonetic accommodation research on native English speakers was extended to non-native speakers and to other languages in work by Kim and colleagues (Kim, Bradlow, and Horton 2007a,b; Kim 2009) using the Wildcat corpus of native and non-native English (Bradlow, Baker, Choi, Kim, and Van Engen 2007). The spontaneous speech in this corpus came from pairs of talkers conversing while playing a picture-matching game. Kim et al. (2007a,b) concentrated on six pairs of female talkers of three types, which differed in terms of the language background of the talkers: (i) native + native, (ii) native + Korean non-native, and (iii) native + Chinese non-native. In a perception experiment using a paradigm similar to that of Pardo (2006), listeners completed an XAB similarity judgment task in which the stimuli comprised 1–1.5 sec samples of speech taken from either the first third or the last third of a conversation. Unlike Pardo (2006), all the stimuli were different in content. The results of this experiment showed both convergence and divergence, depending on the language background of the talker. Native talkers converged to each other, consistent with results found by Pardo (2006). However, they did not converge to non-native talkers, and in one of the native/non-native pairs, the native talker actually diverged from her Korean partner. As for the non-native talkers, one of the Korean talkers was found to converge to her native partner, while one of the Chinese talkers diverged from her native partner. To examine the role of accentedness in patterns of accommodation, the speech of the non-native talkers was also rated for degree of accent by a separate group of native English listeners. These ratings showed that the non-natives with the strongest accents failed to converge to their native partners. On the other hand, the non-native with an intermediate accent converged to her native partner, while the non-native with the weakest accent diverged from her native partner (a result attributed to the possible effects of fatigue arising from the maintenance of a native-like accent).

Kim (2009) expanded this research to include ten additional pairs of talkers: six male-male counterparts to the pair types examined in Kim et al. (2007a,b) and four pairs of native Korean speakers completing the same conversational task in Korean (two female-female and two male-male). Speech samples from the total sixteen pairs were played to native listeners (either English or Korean speakers) for evaluation in the same XAB similarity judgment task and, if applicable, the same accent rating task. The similarity judgment data in this study were largely consistent with those of Kim et al. (2007a,b). Convergence was again found between native speakers, both of English and Korean. However, this convergence was modulated by dialect, such that native talkers of similar dialects converged symmetrically, whereas native talkers of different dialects did not, either converging asymmetrically, diverging, or showing no change. In addition, patterns of accommodation in native/non-native pairs were again affected by the strength of a non-native talker's accent, a finding indicative of a relationship between phonetic accommodation and pronunciation aptitude (see also Lewandowski 2009). Native English speakers were found to diverge from non-native partners with strong accents. In contrast, when the non-native partner had a weak accent, one native English speaker actually converged. Non-native talkers usually converged toward native partners as long as their accent was only intermediate. However, these were in fact all Korean non-natives, as the Chinese non-natives uniformly failed to converge toward their native partner regardless of their degree of accentedness, suggesting that sociocultural factors may play a role in phonetic accommodation as well.

In short, work on L1 speech accommodation has demonstrated that talkers may be significantly influenced in their own native language production by the phonetic details of other talkers' speech. Changes in L1 production occur in talkers engaged in interactive conversation, in talkers asked to shadow another talker, and in talkers simply exposed to a disembodied voice. Although these developments seem to be convergent more often than not, whether they take the form of convergence or divergence appears to be affected by both social and phonological factors. These findings of pervasive accommodation to other talkers' speech are relevant to the study of L2 influence on L1 production because they demonstrate that speech input from others is experienced in an input-output loop not unlike that of auditory feedback from one's own speech—that is, adjustments to L1 production are made not to feedback *per se*, but to speech input generally. Thus, if L2 speech input is experienced in a similar fashion as L1 speech input, then the expectation is that adjustments of L1 production to L2 speech input will occur in an analogous way.

2.2.4 Summary

Evidence from several different research programs converges upon the same finding: L1 production changes in response to the environment. Studies on the Lombard effect in hearing people, degradation and recalibration of production in cochlear implant users, perturbations caused by delayed feedback, sensorimotor adaptation to altered feedback, phonetic convergence in interactive conversation, and spontaneous imitation of model speech all show that speakers make significant adjustments to their L1 production relative to external input, often rapidly so. Given the extent of L1 malleability, it is reasonable to think that L1 might change in response not only to input in L1, but also to input in L2. In the absence of evidence to the contrary, such cross-linguistically influenced changes would then be expected to bear similarities to the patterns of language-internal change observed in speech adaptation to attenuated feedback, to altered feedback, and to auditory input from other talkers. Cross-linguistically influenced changes in L1 production have indeed been found in the study of L2 speech, but before these results are reviewed, it is necessary to provide some background on the tradition of L2 speech in which they were found. This background is provided in the section that follows.

2.3 Modeling the Acquisition of L2 Speech

An abiding concern of psycholinguistic research on speakers of two languages has been the question of whether their languages are represented and processed in one shared system or two separate systems. Since the bilingual lexical/conceptual model of Weinreich (1953), the field has seen the proposal of several other models of bilingualism, mostly of bilingual lexical access and processing—for example, the Word Association Model and Concept Mediation Model (Potter, So, Von Eckardt, and Feldman 1984), the Distributed Conceptual Feature Model (de Groot 1992; van Hell and de Groot 1998), the Revised Hierarchical Model (Kroll and Stewart 1994), and the Inhibitory Control Model (Green 1998). These models are based upon findings from priming and neuropsychological studies, which provide, on the one hand, evidence that use of L1 or L2 often activates the other language even when irrelevant and, on the other hand, evidence that L1 and L2 draw on different neural resources and can be impaired independently of each other (for reviews, see Hartsuiker, Pickering, and Veltkamp 2004; Heredia and Brown 2004; Dong, Gui, and MacWhinney 2005; Kroll and Tokowicz 2005; and Lorenzen and Murray 2008).

The literature on L2 acquisition documents a similar debate regarding the extent to which structures of a developing L2 system are transferred from L1 (cf. "shared") or developed independently from L2 input and a universal, possible innate, linguistic substrate (cf. "separate"). This debate led ultimately to the notion of an "interlanguage" system combining elements of L2, elements of L1, and universal elements manifest in neither language. As such, the study of L2 speech has more often than not proceeded under the working assumption that at least some aspects of L2 are shared with L1. Formalizations of structural sharing between L1 and L2 in the domain of phonology have been attempted by researchers such as Flege and Eefting (1988) and Laeufer (1996) using extensions of the bilingual lexical access models alluded to above, and these are discussed in Section 2.3.4. First, we turn to the three frameworks that have been most influential in the study of non-native and L2 speech perception and production: the Perceptual Assimilation Model (Best 1995), the Speech Learning Model (Flege 1995), and the Perceptual Assimilation Model–L2 (Best and Tyler 2007).

2.3.1 Perceptual Assimilation Model

The Perceptual Assimilation Model (PAM) developed by Best (1993, 1994, 1995) is applicable to the process of L2 phonological acquisition at its very beginning stages. Principally a model of non-native speech perception by naive listeners (i.e., those who have no knowledge of the non-native language), the PAM sets forth a typology of ways in which non-native speech contrasts may be interpreted by naive listeners relative to L1 phonological categories (so-called "perceptual assimilations"). The type of perceptual assimilation that occurs with members of a non-native contrast predicts the degree of difficulty that learners will have with discriminating that contrast: if the members of the contrast are assimilated to different L1 categories, the contrast will be discriminated accurately; if not, the contrast will be discriminated less accurately, to a degree depending upon how equally well the members of the contrast are assimilated to the same L1 category. Thus, non-native contrasts, rather than being uniformly difficult for naive listeners to perceive, are predicted to differ in ease of discriminability according to how they assimilate to L1 categories. With respect to the basis of perceptual assimilation, Best adopts a direct-realist, ecological view of speech perception as based on gestural primitives. According to Best (1994:190), "phonologically mature listeners perceive in nonnative phones information about their gestural similarities to native phonemes", with perceptual assimilation taking place only when a non-native phone and a native phoneme are judged to be sufficiently similar in their articulatory-gestural properties.

Non-native phones that are assimilated to the same L1 category may still be able to be discriminated by the naive listener. This follows from the existence of four basic types of perceptual assimilation: (i) the Two-Category type, in which non-native phones are assimilated to different L1 phonemes; (ii) the Category-Goodness Difference type, in which non-native phones are assimilated to the same L1 phoneme, but one is more different from the L1 phoneme; (iii) the Single-Category type, in which non-native phones are assimilated to the same L1 phoneme with both being equally similar/different from the L1 phoneme; and (iv) the Non-Assimilable type, in which non-native phones are too different from L1 phonemes to be assimilated to any of them and are thus perceived as non-speech sounds. In addition to these types of assimilation, two other types involve perception of a non-native sound as speech, but not as an L1 phoneme (Best 1995): (v) the Uncategorized-Categorized type, in which one non-native sound is assimilated to an L1 phoneme, while the other is perceived as a non-L1 speech sound; and (vi) the Uncategorized-Uncategorized type, in which both non-native sounds are perceived as speech sounds unlike any L1 phoneme.

PAM predicts for the four basic assimilation types that "the discrimination performance pattern for adults should be, from highest performance to lowest, [Two-Category] > ([Non-Assimilable]<=>[Category-Goodness]) > [Single-Category]", although "[Category-Goodness] and [Single-Category] contrasts fall at different ends of a single dimension, in that both involve assimilation of a non-native phone pair to a single native category" (Best 1994:192). Most prior studies of non-native perceptual difficulties in both adults and infants are observed to have focused on Single-Category or Category-Goodness contrasts. While Single-Category contrasts are predicted to be poorly discriminated and Two-Category contrasts easily discriminated, Category-Goodness contrasts are predicted to vary in discriminability according to the phonetic disparity between the two members of the contrast. Similarly, Non-Assimilable contrasts "may vary in degree of discriminability, which will in these cases be determined by variations in salience of the *auditory* differences between pair members...because phonologically sophisticated listeners are expected to perceive them as non-speech sounds" (Best 1994:193). Discrimination of Uncategorized-Uncategorized contrasts is also predicted to vary from poor to good according to the phonetic distance between the non-native sounds and their proximity to L1 categories; on the other hand, discrimination of Uncategorized-Categorized contrasts is generally expected to be good, since in this case the contrast is between an L1-like sound and a clearly non-L1-like sound.

Best (1994) proposes four different hypotheses regarding how 10–12 month-old infants perceive non-native contrasts. The *familiarity hypothesis*, based on the idea that discrimination is based on familiarity, predicts that infants will discriminate all non-native contrasts (with the possible exception of the strong Category-Goodness type) poorly, due to their low degree of familiarity with non-native contrasts. The *strong phonological hypothesis*, on the other hand, predicts that L1-learning infants will be just as biased by L1 phonology in non-native perception as L1 adults are; thus, Two-Category contrasts will be perceived the best, followed by Non-Assimilable and Category-Goodness contrasts, with Single-Category contrasts being perceived the most poorly. The *phonemic contrast hypothesis* predicts that infants will show good discrimination of Two-Category and Non-Assimilable contrasts, but poor discrimination of Category-Goodness and Single-Category contrasts due to "under-differentiated recognition of the coordinated phonetic details within individual native-phone categories" (Best 1994:195). Finally, the *category recognition hypothesis* predicts that infants will show good discrimination of Non-Assimilable and some Category-Goodness and Two-Category contrasts and poor discrimination of Single-Category and other Category-Goodness and Two-Category contrasts because some non-native contrasts of the Category-Goodness and Two-Category varieties are likely to involve gestural coordination patterns difficult to recognize as native patterns.

A large body of research in infant speech perception (e.g., Best, McRoberts, LaFleur, and Silver-Isenstadt 1995; Best and McRoberts 2003) provides the strongest support for the category recognition hypothesis, suggesting that "by at least 10 to 12 months of age, infants have begun to discover the gestural-coordination patterns that identify categories roughly corresponding to phones in their native language" (Best 1994:202), though some categories may still be under-differentiated in comparison to those of adults. Given that they show similar patterns of perceptual assimilation as adults, infants by this age seem to have acquired a considerable foundation of knowledge about the phones of their native language, as "the basis for infants' recognition of the language-specific properties of native and nonnative phones is the detection of evidence about the constellation of coordinated articulatory gestures that are associated with specific phones in the native language" (Best 1994:203).

Thus, both adults and infants have been shown to perceptually assimilate non-native phones to native phonemes, and evidence has been found for all four basic types of perceptual assimilation described by the PAM. Data in support of the Non-Assimilable type of perceptual assimilation were collected by Best, McRoberts, and Sithole (1988) in an experiment testing the perception by English listeners of the contrast among the various Zulu clicks. The results of three experiments using discrimination tasks showed that both mature adults and young infants from 6 to 14 months of age discriminated the clicks very well, in contrast to the patterns of poor discrimination performance for other, more familiar contrasts consistently reported in other studies. In one of the experiments, adults continued to show very good discrimination even when the greatest acoustic difference between click categories—namely, a difference in amplitude—was eliminated from the stimuli. That accuracy of click discrimination was a function of L1 background, rather than due to some inherent discriminability of clicks, was demonstrated in a later study by Best, Traill, Carter, Harrison, and Faber (2003), who found that L1 English speakers discriminated certain !Xóõ click contrasts significantly better than L1 speakers of other African click languages. Counter to what one might expect—that experience with clicks would make click-language speakers better able to discriminate unfamiliar clicks—they performed worse overall than the L1 English speakers. This result follows from the fact that they, unlike the L1 English speakers, had L1 click phonemes that could serve as perceptual attractors for both non-native clicks of a minimal contrast. The conclusion was thus that "phonemic perception entails assimilation of nonnative speech sounds to native categories whenever possible", but that "when they are not assimilated, perception focuses either on purely auditory or phonetic (articulatory) properties" (Best et al. 1988:358).

Best, McRoberts, and Goodell (2001) conducted another study to test for the other three types of perceptual assimilation predicted by the PAM, based on the observation that there is variation in the discrimination of non-native speech contrasts by adults, which "does not depend on the presence or absence of the critical phonetic/acoustic features in native speech" (Best et al. 2001:776). The predictions of the PAM were first compared to those of the Speech Learning Model and the Native Language Magnet Theory (Kuhl and Iverson 1995), with the conclusion that "only PAM makes explicit predictions about assimilation and discrimination differences for diverse types of non-native contrasts"—namely, that "non-native speech perception is strongly affected by listeners' knowledge (whether implicit or explicit) of native phonological equivalence classes", and that "listeners perceptually assimilate non-native phones to native phonemes whenever possible, based on detection of commonalities in the articulators, constriction locations and/or constriction degrees used" (Best et al. 2001:777).

Best et al. (2001) ran two experiments to test these predictions. In Experiment 1, English speakers were tested on their perception of three types of non-native contrasts from Zulu: (i) the contrast between voiced and voiceless lateral fricatives, a Two-Category contrast with respect to English; (ii) the contrast between voiceless aspirated and ejective velar stops, a Category-Goodness contrast with respect to English; and (iii) the contrast between plosive and implosive voiced bilabial stops, a Single-Category contrast with respect to English. Measurements of the stimuli on several acoustic properties showed similarities among the three contrasts in terms of the number of acoustic differences between the members of the contrast. Despite these similarities, however, the three contrasts were discriminated differently, with Two-Category > Category-Goodness > Single-Category, as predicted. Furthermore, it was found that the performance of subjects who assimilated the three contrasts in the expected manner improved when the first member of the contrast was made more similar to an English phone, suggesting that listeners were sensitive to within-category variation. In addition, the particular assimilations, as measured by subjects' spelling of the sounds, generally seemed to be based on the stimuli's articulatory properties, consistent with "the direct realist position that listeners detect information in signals about the nature of the event that produced the signal" (Best et al. 2001:786). Finally, recency effects were found in the case of the Single-Category contrast, which, when compared to similar recency effects in the click discrimination experiment of Best et al. (1988), "appear to be associated specifically with detection of nonlinguistic as opposed to phonological or phonetic differences" (Best et al. 1988:786). Experiment 2 tested perception of the Tigrinya contrast between bilabial and alveolar ejectives, another Two-Category contrast with respect to English. These phones were found to be discriminated at a high level similar to that of the Two-Category contrast in the Zulu perception experiment, demonstrating that "the [Two-Category] assimilation pattern...applies to constriction location contrasts as well as laryngeal gesture contrasts" (Best et al. 2001:789). The fact that the Two-Category contrasts in both experiments showed discrimination performance significantly below that of native English contrasts was interpreted as evidence that listeners were influenced both by phonologically relevant information in the signal used to perceive a native phonological contrast, as well as by phonologically irrelevant differences between native phonemes and similar non-native phones.

Following from these results, Best et al. (2001:786) noted that "detection of contrastive phonological distinctions versus non-contrastive phonetic details versus nonlinguistic auditory properties is somehow differentiated in non-native speech perception". They concluded that "familiarity with the typical phonetic form of native consonants aids rather than hinders discrimination [of nonnative consonants], whether the listener is attending for information about phonological contrast, or phonetic goodness of fit to a single phoneme, or nonlinguistic stimulus variations", insofar as "native speech experience results in more stable perception of tokens that are more nativelike" (Best et al. 2001:790). They also suggested that "infants progress developmentally from detection of only nonlinguistic...information in speech, to recognition of how phonetic variants fit into...language-specific phonetic classes, to eventually discovering the phonologically contrastive functions those phonetic classes serve in distinguishing native words" (Best et al. 2001:791).

The PAM claim that between-category relationships in non-native speech perception are established on the basis of phonetic similarity was also supported by the findings of Hallé, Best, and Levitt (1999), who argued that "articulatory-phonetic considerations, and not only abstract phonological considerations and descriptions, must be taken into account to explain non-native speech perception performance" (Hallé et al. 1999:303). In this study, two perception experiments were conducted involving identification and discrimination of American English approximants. Results of these experiments indicated that French listeners had trouble with /r/, assimilating it to /w/, even though they had an /r/-/l/ contrast in their L1. These results are unexpected if one only considers the fact that French has a rhotic-lateral contrast, but are unsurprising if one considers the phonetic differences between the two languages' rhotic phonemes. French speakers were also found to discriminate /w/-/j/ better than American English or Japanese speakers, even though all these groups had this contrast in their L1. This finding was attributed to the more densely packed inventory of glides in French (also containing a labial-palatal glide /ɥ/), which may have resulted in greater perceptual sensitivity to contrasts between glides in general. Although Hallé et al. were careful not to rule out the role played by L1 phonology in non-native speech perception, they concluded that "the comparative articulatory-phonetic details of two or more languages are more clearly associated with detailed variations in performance on different non-native contrasts than are abstract phonological descriptions alone" (Hallé et al. 1999:303). In other words, non-native speech perception is governed more closely by proximity in phonetic realizations than by parallelism in phonemic representations.

Whereas the above studies delved into the question of perceptual assimilation of non-native segments in the same, unmarked context, other studies have focused on perceptual assimilation of non-native phonotactics. In a classic study by Dupoux, Hirose, Kakehi, Pallier, and Mehler (1999), the results of two experiments were reported which suggested that non-native speech perception is affected not only by L1 inventory constraints, but also by L1 phonotactic constraints. In Experiment 1, Japanese speakers were found to perceive "'illusory' vowels" within consonant clusters—specifically, the vowel [u] between the two consonants of VCCV stimuli—while in Experiment 2, they were found to have trouble discriminating between VCCV and VCuCV stimuli. French speakers, in contrast, did not show this perceptual behavior; however, in comparison to Japanese speakers, they had trouble distinguishing between stimuli differing in vowel length, which is distinctive in Japanese but not in French. Thus, it was argued that "models of speech perception have to be revised to account for phonotactically based assimilation" (Dupoux et al. 1999:1568).

Like Dupoux et al. (1999), Hallé, Segui, Frauenfelder, and Meunier (1998) provided experimental results showing a strong influence of L1 phonotactics on non-native speech perception. Several experiments were conducted with French listeners using a variety of tasks and methodologies that showed that illegal clusters */dl, tl/ were perceptually assimilated to legal clusters /gl, kl/. First, speakers identified the illegal dental clusters as the legal velar ones in an open response task. Second, in a phonemic gating task, they initially identified dentals in early gates, but then shifted to identifying velars in later gates as /l/ became audible, demonstrating that the percept of the legal clusters was not due to any defect of the dental portion of the stimuli. Finally, in a phoneme monitoring task, speakers were slower to identify phonemes when they occurred in illegal clusters. Thus, these results, which showed a phenomenon the authors dubbed "contextual perceptual assimilation", suggested that non-native speech perception involves "automatic integration of low-level phonetic information into a more abstract code determined by the native phonological system" (Hallé et al. 1998:592).

Hallé and Best (2007) followed up on the findings of Hallé et al. (1998) by extending the basic experiment to two other groups (Hebrew speakers and American English speakers) in order to examine whether the perceptual difficulties of French speakers with */dl, tl/ found by Hallé et al. (1998) were due specifically to French phonotactic restrictions or to other factors like language-independent perceptual difficulty associated with these clusters. Discrimination and categorization experiments showed, first, that Hebrew speakers, whose L1 permits all of /dl, tl, gl, kl/, did not have this sort of difficulty with the dental clusters, and second, that American English speakers, whose L1 realizes the voiced-voiceless stop distinction differently from French, patterned with the French speakers, though they had greater difficulty with the /dl/-/gl/ distinction. These results were interpreted as evidence for "language-specific phonotactic perceptual assimilation, with modest contributions from language-specific phonetic settings" (Hallé and Best 2007:2899).

In sum, the PAM claims that perceptual assimilation of non-native phones to native L1 categories falls into one of six types (Two-Categories, Category-Goodness, Single-Category, Uncategorized-Categorized, Uncategorized-Uncategorized, and Non-Assimilable) that differ predictably in terms of discrimination difficulty. Perceptual assimilation is based upon comparisons of phonetic—specifically, articulatory—details of native and non-native sounds and is, moreover, highly influenced by phonological context. Although the PAM is, strictly speaking, a model of non-native perception by naive listeners (i.e., not L2 learners), implications for the perception of L2 contrasts by L2 learners follow quite naturally from the perceptual claims of the model; however, the connection to L2 perception is only addressed explicitly in a later version of the model, which is not extended to account for L2 production (see Section 2.3.3). Consequently, in the domain of L2 production a different model (which does make an explicit connection between L2 perception and L2 production) has proven to be more influential—namely, the Speech Learning Model.

2.3.2 Speech Learning Model

Whereas the PAM provides an account of non-native speech perception, the Speech Learning Model (SLM) developed by Flege (1988, 1992, 1995) provides an account of L2 speech learning based upon the idea that "phonetic systems reorganize in response to sounds encountered in an L2 through the addition of new phonetic categories, or through the modification of old ones" (Flege 1995:233). Contrary to the notion of a critical period, the SLM postulates that learning mechanisms used in L1 acquisition are available throughout life (see also Wode 1994), and that an L1 phonetic category encoding language-specific features of an L1 sound continues to develop in adulthood under the influence of all sounds (regardless of language) identified with that category. Furthermore, phonetic categories of L1 and L2 are posited to exist in a shared system, where there is a general pressure to keep them distinct.

From these postulates follow seven hypotheses (Flege 1995:239), which can be condensed into four. First, it is hypothesized that "sounds in the L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level". Second, a new phonetic category may be formed for an L2 sound given sufficient dissimilarity from the closest L1 sound, although this becomes increasingly unlikely at older ages of learning; if a new phonetic category is formed for an L2 sound, however, and it contains the same information as would be found in a native speaker's, then the L2 sound will be able to produced accurately. Third, it is predicted that "category formation for an L2 sound may be blocked by the mechanism of equivalence classification", such that "a single phonetic category will be used to process perceptually linked L1 and L2 sounds", which will eventually come to "resemble one other in production". Fourth, it is expected that "the phonetic category established for L2 sounds by a bilingual may differ from a monolingual's" in two scenarios: (i) when "the bilingual's category is 'deflected' away from an L1 category to maintain phonetic contrast", or (ii) when "the bilingual's representation is based on different features, or feature weights".

Thus, at its most basic level the SLM asserts continuity between L1 speech learning mechanisms and L2 speech learning mechanisms, with the crucial difference between L1 acquisition and L2 acquisition being one of prior experience. Central to the model is its account of inaccurate production of an L2 sound in terms of identification with a similar L1 category via EQUIVALENCE CLASSIFICATION. Equivalence classification, "a basic 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" (Flege 1987b:49), is responsible for speech communication in general being able to occur successfully in spite of intra- and inter-speaker phonetic variability. Applied to the learning of L2 phones, equivalence classification is argued to target L2 phones that are "similar" to L1 phones, but not so-called "new" phones. The equivalence classification of "similar" L1 and L2 phones ultimately limits the accuracy with which the L2 phone can be produced because the L2 phone is influenced by the properties of the similar, but non-identical L1 phone. As Flege (1987b:62) puts it, "if equivalence classification prevents L2 learners from developing a separate phonetic category for similar L2 phones, they may be unable to produce similar phones in L2 and L1 authentically because they need to implement [the phones] in both L2 and L1 using the same phonetic category"; consequently, "the 'merging' of the phonetic properties of similar L1 and L2 phones" in one category may result in "an upper limit on phonetic approximation for similar L2 phones".

The distinction among "new", "similar", and "identical" L2 phones is critical for the formulation of predictions regarding accuracy of L2 production under the SLM. Different degrees of similarity between L1 and L2 sounds have been discussed by a number of L2 researchers (e.g., Brière 1966; Wode 1978), but the SLM attempts to formalize these notions of L1-L2 similarity in the face of "the lack of an objective means for gauging degree of perceived cross-language phonetic distance" (Flege 1995:264).[1] According to the SLM, "similar" L2 phones "differ systematically from an easily identifiable counterpart in L1", whereas "new" L2 phones "have no counterpart in L1" (Flege 1987b:48). "Identical" L2 phones, like "similar" phones, have a clear L1 counterpart, but are so close to the nearest L1 sound that the difference between the two is negligible. The classification of L2 sounds into these three types is based on transcription practices, acoustic proximity, and perceptual similarity (Flege 1996:16–18). Thus, an L2 sound that is "identical" to an L1 sound is typically transcribed with the same IPA symbol, is not significantly different acoustically, and is not perceptibly different. "Identical" L2 sounds undergo equivalence classification with L1 sounds in the same way that "similar" L2 sounds do, but since they are imperceptibly different from L1 sounds, this does not have a negative impact on their production. An L2 sound that is "similar" to an L1 sound is also transcribed with the same IPA symbol, but shows significant acoustic differences from the L1 sound that are perceptible to listeners. A "new" L2 sound is different in all of the aforementioned ways: it is transcribed differently from any sound of the L1 and is both acoustically and perceptually distinct from the nearest L1 sound. It should be noted that in classifying an L2 sound as "identical", "similar", or "new", consideration of transcriptions is supposed to serve as a "preliminary step", since, as acknowledged by Flege (1996), transcription practices and conventions vary widely, rendering it "necessary to supplement the phonetic symbol test with additional acoustic criteria" as well as with perceptual data and other processing measures.

Although the SLM provides three clear criteria for the classification of L2 sounds as "identical", "similar", or "new", it is not fully explained how these criteria interact with each other—in particular, how they are to be reconciled when they conflict. For example, transcription conventions based partly on phonemic considerations may often be at odds with acoustic phonetic comparisons. Since in the SLM L1 and L2 sounds are supposed to be related at a position-sensitive allophonic level rather than at a phonemic level, it is reasonable to think that in these cases the acoustic phonetic comparisons should prevail. However, this interpretation is inconsistent with the way the "new" vs. "similar" distinction is applied throughout the literature, including Flege and Hillenbrand's (1984) study of American English speakers' production of French /u/ and French /y/ (in the words *tous* /tu/ 'all' and *tu* /ty/ 'you'). For L1 English speakers, L2 French /u/ is considered to be a "similar" vowel with an L1 counterpart in English /u/, while L2 French /y/ is considered to be a "new" vowel

[1] Flege (1995:264) observes that cross-language phonetic distance "might be assessed in terms of the sensory (auditory, visual) properties associated with L1 and L2 sounds" or "differences in perceived gestures". The SLM is, therefore, amenable to a gestural framework like that of Best (1995). Nonetheless, specific criteria for determining cross-language similarity are described in terms of acoustic rather than gestural comparisons (e.g., Flege 1996). As such, work in the SLM has generally been conducted within an acoustic framework.

with no L1 counterpart. This classification contrasts with the acoustic facts, which show that in the given context French /y/ is actually more similar than French /u/ to English /u/, which is significantly fronted in the context of alveolars (Strange, Weber, Levy, Shafiro, Hisagi, and Nishi 2007; Levy 2009). Moreover, documented perceptual assimilation patterns show that English speakers tend to identify French /y/, as well as German /y/, as close to English /u/ (Strange, Levy, and Lehnholf, Jr. 2004; Polka and Bohn 1996). In the end, though, French /u/ and /y/ are argued to be, respectively, "similar" and "new" vowels for L1 English speakers, and the patterns in English speakers' production of these sounds are consistent with these classifications. These findings suggest that in the calculation of cross-language similarity for L2 learners, phonemic proximity (as encoded in transcription practices) may trump phonetic proximity, but this point is not addressed explicitly in the SLM. The issue of resolving conflicting sources of information regarding cross-language similarity is discussed further in Chapter 6.

However these conflicts are resolved, if an L2 phone is deemed "similar" or "identical" to an L1 sound, it will undergo equivalence classification with the L1 category, and this will limit the accuracy with which a "similar" L2 sound can be produced because it will be influenced by disparate features of the L1 category. The insight behind equivalence classification is that it provides an explanatory account of the apparent effects of age on L2 phonological acquisition. Flege (1987b:50) observes that "if humans rely increasingly less on sensory information in making categorical decisions as they mature, and if, at the same time, they become capable of identifying an increasingly wider range of phones as belonging to a phonetic category, it may become increasingly difficult for L2 learners to note the phonetic (but not auditory) difference between 'similar' phones in L1 and L2". To put it another way, it is the consequences of linguistic experience, not neurological developments *per se*, that result in the longitudinal decline in ability to acquire L2 sounds like a native speaker. A "new" L2 sound, by virtue of being significantly different from any previously experienced L1 sound, is not analogized to an L1 sound and causes the formation of a new phonetic category. In contrast, a "similar" L2 sound, by virtue of being similar in one or more ways to a previously experienced L1 sound, is analogized to this L1 sound.[2] Given the experiential basis of equivalence classification, it occurs with increasing probability as age of L2 learning increases, resulting in the perceptual linkage of close L1 and L2 sounds. All is not lost when L2 sounds are initially identified with L1 sounds, however. Consistent with the tenet of continuous development of the phonetic system, the SLM claims that "as L2 learners gain experience in the L2, they may gradually discern the phonetic difference between certain L2 sounds and the closest L1 sound(s)" (Flege 1995:263). Discernment of these differences may

[2]The amount of dissimilarity required for an L2 sound to count as "new" vs. "similar" is still not well understood. Flege (1996:42), for instance, observes that "it remains to be determined if there is a phonetic difference threshold that, once crossed, triggers the formation of a new phonetic category".

turn a "similar" L2 sound into a "new" L2 sound, allowing an independent phonetic category to be established for the L2 sound.

A major way in which the SLM differs from the PAM—and the principal reason the SLM is more relevant to the present study—is that the SLM overtly addresses the connection between L2 perception and L2 production. Because "production of a sound eventually corresponds to the properties represented in its phonetic category representation" (Flege 1995:239), an L2 sound linked to a new, unique phonetic category resembling a native speaker's is predicted to be produced accurately, while L1 and L2 sounds perceptually linked to the same category, so-called "diaphones" (Weinreich 1957), are predicted over time to approximate each other in production. At the same time, however, following from the notion of L1 and L2 sounds existing in the same phonological space, the SLM allows for the possibility of an L2 phonetic category dissimilating from an L1 category for the sake of cross-linguistic contrast. In this way, L2 speech acquisition may lead to either convergence or divergence between L1 and L2 sounds.

It should be noted that although the SLM's notion of perceptual linkage of similar L1 and L2 sounds has often been used to account for how L1 is transferred to L2 production, it provides at the same time a theoretical formulation of how L2 can be transferred to L1 production. In fact, the prediction of L1 phonetic drift resulting from L2 learning is made explicit in the model:

> ...cross-language phonetic interference is bidirectional in nature. The [Speech Learning Model] predicts two different effects of L2 learning on the production of sounds in an L1, depending upon whether or not a new category has been established for an L2 sound in the same portion of phonological space as an L1 sound. (Flege 1995:241)

Specifically, if a new category is not established for an L2 sound (i.e., the L2 sound is linked to an L1 category), then L2-to-L1 interference is predicted to be convergent in nature. On the other hand, if a new category is established for the L2 sound, then L2-to-L1 interference (if it occurs; see Flege 2002) is predicted to be divergent in nature, motivated by the pressure to maintain cross-linguistic contrast between L1 and L2 sounds in a shared phonological space.

Given that the SLM is specifically concerned with "ultimate attainment of L2 pronunciation", most work in this model (e.g., Flege 1991b; Flege, Munro, and MacKay 1995a; Flege, Bohn, and Jang 1997; MacKay, Flege, Piske, and Schirru 2001; Flege et al. 2003) has focused on "bilinguals who have spoken their L2 for many years, not beginners" (Flege 1995:238). This work includes studies of the production and perception of vowels, initial consonants, and final consonants and is largely found to provide support for the hypotheses of the SLM.

First, the hypothesis that L1 and L2 sounds are related at a position-sensitive allophonic level is supported by evidence from L1 Japanese speakers' acquisition of

L2 English /r/ and /l/. Japanese learners of English are often observed to have trouble with this contrast as a result of having only one liquid phoneme in their L1 (e.g., Yamada and Tohkura 1992; Yamada 1995). However, the degree of success they have in perceiving the contrast has been shown to vary, depending on factors such as word familiarity (Yamada, Tohkura, and Kobayashi 1996). One variable that has been shown to affect both perception and production of this contrast is phonological environment. Sheldon and Strange (1982), for example, found that Japanese learners' perception of the English liquid contrast was poor within prevocalic consonant clusters, but good word-finally. Similarly, Strange (1992) found that Japanese learners were generally better at perception and production of word-final liquid contrast than word-initial liquid contrast. This asymmetry between word-initial position and word-final position was attributed by Sheldon and Strange (1982) to more robust acoustic differences between the two liquids in word-final position, which may also account for their disparate adaptations in loanwords in only this position. While English /r/ and /l/ do not differ in the way they are adapted into Japanese in initial position, they do differ in final position (/r#/ > /a/, /l#/ > /ru/), suggesting that a position-sensitive relation to L1 sounds may be responsible for the more accurate perception and production of the contrast in final position.

Second, the hypothesis that a novel L2 sound may result in the formation of a new phonetic category and be produced accurately is supported by evidence from several production studies. Catford and Pisoni (1970) demonstrated that phonetically untrained English speakers could learn to accurately produce novel consonants and vowels in isolation, and that articulatory training consistently resulted in better performance than auditory training. In a more linguistic task involving words, Flege (1987b) found that L1 English speakers experienced in French produced the L2 French vowel /y/, a "new" vowel from the perspective of English, relatively accurately; in fact, they did not differ from native French speakers in this regard. Chang, Haynes, Yao, and Rhodes (2010) replicated this result with L1 English learners of Mandarin, who did not differ significantly from native Mandarin speakers in their production of Mandarin /y/. Focusing on the "new" English vowel /æ/, Flege (1996) found that L1 Dutch speakers proficient in English produced /æ/ accurately, and this result was replicated for L1 German speakers by Bohn and Flege (1996). This hypothesis also predicts that the likelihood of formation of a new phonetic category will decline with age. This is consistent with a myriad of findings in the literature documenting declines in individual category production and overall production accuracy with increasing age of L2 learning (e.g., Flege 1991a; Flege et al. 1995a,b; Flege, MacKay, and Meador 1999; Flege, Yeni-Komshian, and Liu 1999; Flege et al. 2003), although sometimes the age-correlated differences are attributed to differences in input to early and late learners (e.g., MacKay et al. 2001).

Third, the hypothesis that a novel L2 sound may instead be identified with a similar L1 sound and be produced inaccurately is supported by numerous studies of consonant and vowel production. In Flege's (1987b) study, for example, American

English speakers were found to produce French /u/ with F_2 values that were too high, under influence from the high-F_2 /u/ of English. These results are consistent with those of Chang et al. (2010), who found that American English-speaking learners of Mandarin produced Mandarin /u/ with F_2 values higher than native Mandarin speakers'. As for consonants, the English speakers in Flege (1987b) were also found to produce French voiceless stops with VOTs that were too long, under influence from the long-lag VOT of English voiceless stops. Several other studies have documented the difficulties that speakers of languages with short-lag voiceless stops have with learning the long-lag voiceless stops of English; typically they are found to produce the English stops with VOTs that are too short, under influence from their L1 voiceless stops (e.g., L1 Arabic: Flege 1980; L1 Dutch: Flege and Eefting 1987a; L1 Spanish: Flege 1991a). As discussed above, the notion of perceptual linkage between L1 and L2 sounds also predicts influence of the L2 sound on the L1 sound. There are findings that bear out this prediction as well, and these are discussed in Section 2.4.3.

Finally, the hypothesis that a bilingual's L2 categories may differ fundamentally from a monolingual's is supported by data from several studies of L2 English demonstrating that "certain vowel errors persist in the speech of highly experienced L2 learners", likely due to "use of non-English feature specifications" (Flege 1995:253). For example, in a study of English vowel production by L1 Arabic learners living in the U.S. (Munro 1993), perceptual data showed that vowels of even the most experienced learners, who had been living in the U.S. for over 15 years, were perceived as accented. At least some of this accentedness seemed to be attributable to non-native-like production of duration differences between tense and lax English vowels. These differences were found to be exaggerated by L1 Arabic learners, suggesting that the L2 tense and lax categories might have been interpreted not as tense and lax categories, but as long and short categories (which are found in the L1). The use of different feature specifications in a bilingual's L2 categories was also likely responsible for the persistence of accented vowels in the speech of L1 Italian learners of English (Munro, Flege, and MacKay 1996). Here it was found that learners' productions of English /ɚ/ showed a particularly pronounced discrepancy between identification data and accent rating data. Productions of /ɚ/ by subjects who began learning English after the age of ten were identified well, but nevertheless rated as accented—a result that may have been due to failure to use "the retroflex feature (i.e., energy in the region of F_3) which is used to distinguish /ɚ/ from other English vowels...but apparently is not used in Italian" (Flege 1995:254).

This hypothesis also predicts differences between a bilingual's L2 category and a monolingual's due to dissimilation from the bilingual's L1 categories, and evidence of such dissimilation is found in the literature (see Laeufer 1996). The SLM does not make an explicit statement regarding the effect of age on dissimilation, but from the hypothesis that new phonetic categories are more often formed at earlier ages of L2 learning, it follows that this sort of dissimilation should be found more often in early as opposed to late bilinguals, since early bilinguals are more likely to have separate phonetic categories for similar L1 and L2 sounds that can dissimilate from each other. Indeed, this is the pattern that is found in previous studies. In Flege et al. (2003), for example, while late L1 Italian-L2 English bilinguals were found to produce the English vowel /e/ (realized as [eɪ]) as more monophthongal than monolingual native English speakers (under influence from the Italian monophthong /e/), early bilinguals were found to produce English /e/ as more diphthongized than monolingual native English speakers, suggesting that their L2 /e/ had dissimilated from the corresponding L1 /e/. This sort of divergence between L1 and L2 categories can also affect the realization of L1 categories, consistent with the assumption of bidirectionality of cross-linguistic influence (Flege 2002).

To summarize, by claiming that speech learning continues throughout life with much the same learning mechanisms, the SLM attributes differences between L1 and L2 speech learning in large part to one's prior linguistic experience. Age, as a proxy for experience, is therefore expected to have a significant effect on learning outcomes. L1 experience influences the way L2 sounds are perceived in that an L2 phone "similar" to a previously experienced L1 category will tend to undergo equivalence classification with it, leading to the perceptual linkage of the two sounds to the same category and thereby limiting the accuracy with which the L2 phone can be produced. Even "new" L2 sounds that induce the formation of their own phonetic category may be influenced by L1 sounds, as dissimilation may occur between L1 and L2 phonetic categories so as to maximize contrast within a shared phonological system. The SLM, however, assumes that cross-language influence is mutual. Thus, convergence between L1 and L2 phones linked to the same category is predicted to affect the production of the L1 sound as well; likewise, divergence between L1 and L2 categories is also expected to affect the L1 sound. These elements of the SLM are straightforward, but the model's predictions are dependent on relationships of cross-linguistic similarity, the determination of which was noted to be ambiguous in some respects under this model. This issue is taken up in an updated version of the PAM, the Perceptual Assimilation Model–L2.

2.3.3 Perceptual Assimilation Model–L2

In a revised version of the Perceptual Assimilation Model, the Perceptual Assimilation Model–L2 (PAM-L2, Best and Tyler 2007), the principles of the original PAM for non-native speech perception are extended to L2 speech perception. While the PAM's metatheoretical assumption of a gestural basis for speech perception remains the same in the PAM-L2, the PAM-L2 expands upon the PAM by incorporating the influence of an L2 learner's developing phonetic and phonological knowledge of L2, thus allowing for perceptual assimilation at the gestural, phonetic, and phonological levels. The novel possibility of assimilation at the phonological level is the feature of this model that most differentiates it from the SLM. In fact, the PAM-L2 claims that "contrasts at the functional linguistic level of the L1 phonology and their relationship to phonological contrasts of the L2 are as important to perceptual learning as phonetic categories in the two languages" (Best and Tyler 2007:26).

Like the PAM, the PAM-L2 makes its predictions based on "the ecological, direct-realist premise that the focus of speech perception is on information about the distal articulatory events that produced the speech signal" (Best and Tyler 2007:22). Gestural dimensions, not acoustic properties, are the phonetic primitives assumed to form the basis for judgments of cross-linguistic similarity and, thus, perceptual assimilation. It bears repeating that this is a radical difference from the SLM, which assumes no such gestural framework. Since the PAM-L2 claims that "language users become special purpose devices for perceiving transformational invariants in the locations, amplitudes and phasings of speakers' vocal tract gestures in the L1 and/or L2", it follows that, contrary to the SLM's emphasis on the phonetic category, "mental representations of phonetic categories are not required for L2 perceptual learning" (Best and Tyler 2007:25).

Best and Tyler (2007) introduce the PAM-L2 by first exploring the commonalities between the findings of work in the PAM on non-native perception by naive listeners and the findings of work in the SLM on L2 perception and production by relatively experienced L2 learners. They observe that although many principles and predictions of the PAM and the SLM converge, the models describe different populations and are thus not interchangeable. Specifically, the PAM focuses on naive perceivers of non-native speech, functional monolinguals who are "not actively learning or using an L2, and are linguistically naive to the target language" (Best and Tyler 2007:16); these individuals may have "relatively passive exposure to a language other than the L1, that is, for which the listener has made little or no active attempt to learn the language" and/or "limited L2 instruction, especially classroom-only instruction with instructors who have a strong L1 accent" (Best and Tyler 2007:34). In contrast, the SLM focuses on L2 learners, individuals who can be said to be "in the process of *actively learning* an L2 to achieve functional, communicative goals, that is, not merely in a classroom for satisfaction of educational requirements" (Best and Tyler 2007:16); in particular, work in the SLM has generally examined experienced L2 learners.

The distinguishing feature of the PAM-L2 with respect to the PAM is the claim that "the phonological level is central to the perception of L2 speech by [L2] learners" (Best and Tyler 2007:23). Given that naive perceivers have no knowledge of the non-native language they are hearing (including its phonology), there is no place for phonological knowledge of the non-native language in the PAM. On the other hand, in the PAM-L2, L2 learners' knowledge of the L2 on both a phonetic level and a phonological level plays a role in how sounds are related to one another across languages. In this respect, the model attempts to address the question raised in Section 2.3.2 regarding the determination of cross-language similarity, an issue that "has not yet received adequate treatment in any model of nonnative or L2 speech perception: How listeners identify nonnative phones as equivalent to L1 phones, and the level(s) at which this occurs" (Best and Tyler 2007:26).

The PAM-L2 posits that similarity may be perceived at the gestural, phonetic, or phonological level, but that “the perceptual objects/events of interest depend on the perceiver’s perceptual goals or focus of attention” (Best and Tyler 2007:25). Thus, the use of gestural, phonetic, and/or phonological information is related to the L2 learner’s level of analysis of the signal, which may vary according to the stage in acquisition (e.g., practicing individual sounds vs. acquiring new words). In particular, it is predicted that expansion of the lexicon, which usually occurs earlier in L2 acquisition than in L1 acquisition, “is likely to exert forceful linguistic pressure for the L2 learner to ‘re-phonologize’ perception of the target contrasts, whereas the naive listener has no such motivation” (Best and Tyler 2007:32). It follows that classroom L2 learners, who typically gain an almost immediate awareness of which L2 sounds are distinctive in the language (an awareness reinforced by simultaneous learning of orthography along with lexical items), might be expected to show an especially strong tendency to “tune” perception early on in favor of successfully distinguishing the meaningful contrasts of the L2.

Although any of the gestural, phonetic, and phonological levels may play a larger role than the other two levels at a given point in L2 acquisition, the PAM-L2 adopts a broad view of the way they may interact, stating that “L1-L2 differences at a gestural, phonetic, or phonological level may each influence the L2 learner’s discrimination abilities, separately or together, depending on the context or the perceiver’s goals” (Best and Tyler 2007:25). The model focuses in particular on cross-linguistic similarity at the phonetic and phonological levels, and in this respect the PAM-L2 departs from the SLM, which adheres to similarity relations between L1 and L2 sounds at the phonetic level only. The notion of similarity at the phonological level—indeed, the notion of a phoneme—is dependent upon knowledge of the lexicon, since the learner must know words in order to know which sounds can distinguish words. Consequently, the phonological level in this model is essentially a “lexical-functional level”, where similarity of an L2 phonological category to an L1 phonological category entails that “the phonological category has a similar contrastive relationship to surrounding categories in the phonological space”, although similarity between two categories at the phonological level “does *not* automatically imply equivalence or even perceived similarity at the phonetic level” (Best and Tyler 2007:27–28). Given a close enough degree of phonological similarity between L1 and L2 sounds, perceptual assimilation may occur at the phonological level since “listeners may identify L1 and L2 sounds as functionally equivalent (assimilated phonologically)” (Best and Tyler 2007:26).

The example used to illustrate assimilation at the phonological versus phonetic level is that of the rhotic /r/, which corresponds to very different phonetic realizations across languages such as French and English. While /r/ in European French is produced as a uvular trill [ʀ] or fricative [ʁ], /r/ in American English is produced as an alveolar approximant [ɹ]. Despite the large phonetic disparity between these two rhotics, L1 American English speakers learning European French are often observed to equate them (e.g., producing [ɹ] for French /r/).[3] This sort of evidence suggests that cross-language identification of L1 and L2 sounds may occur at the phonological level, even if the sounds are clearly distinguishable at the phonetic level. However, such phonological equation does not entail phonetic equation. English-speaking learners of French do not all produce French /r/ as [ɹ]; moreover, the ones who do produce [ɹ] in French do not necessarily do so all the time. Thus, L1 and L2 categories equated at a higher level may nevertheless be dissociated at a lower level.

The PAM-L2's predictions for L2 perception are elaborated for five possible alignments of L1 and L2 sounds (the L2 sounds being conceptualized in terms of pairs of sounds that minimally contrast, in keeping with the PAM). The first possible alignment is one in which the two L2 phonemes are not assimilated to the same L1 phoneme: one L2 phoneme is perceptually assimilated to an L1 phoneme, while the other L2 phoneme is not (i.e., Two-Category or Uncategorized-Categorized assimilation in the PAM). In this case, the assimilated L2 sound may be assimilated at the phonetic level and/or the phonological level. If the L2 sound is similar to the L1 phoneme on the phonetic level, then it will be perceived as a good exemplar of the L1 category, and perceptual assimilation will take place on both the phonetic level and the phonological level. On the other hand, if the L2 sound is dissimilar from the L1 phoneme on the phonetic level, then it will be perceived as a deviant exemplar of the L1 category, and perceptual assimilation will take place only on the phonological level. In either case, because the other member of the L2 contrast is either assimilated to a different L1 category or not assimilated to any L1 category, discrimination of the L2 contrast is predicted to be good. For example, L1 English learners of French and L1 German learners of French are both expected to have little trouble distinguishing French /r/ and /l/. On the phonological level, both groups of learners are expected to assimilate L2 French /r/ to their L1 /r/. On the phonetic level, L1 German learners are much more likely than L1 English learners to assimilate L2 French /r/ to their L1 /r/ (which is realized as uvular [ʀ, ʁ] in standard German as well). However, this does not change the fact that neither group is likely to also assimilate French /l/ to their L1 /r/ (probably they will assimilate it to their L1 /l/). Consequently, discrimination of French /r/ and /l/ is predicted to be good for both groups.

The second possible alignment of L1 and L2 sounds is one in which the two L2 phonemes are perceptually assimilated to the same L1 phoneme, but with different goodness of fit (i.e., Category-Goodness Difference assimilation in the PAM). In this case, both L2 phonemes are assimilated to the L1 phoneme on the phonological level, but only the phonetically more similar L2 phoneme is assimilated to the L1 phoneme on the phonetic level; the phonetically less similar L2 phoneme

[3] The orthographic equivalence of French /r/ and English /r/ may play a role in this cross-language identification; however, it is noted that "the orthographic commonality may itself reflect the similar patterning of rhotics across the two languages in terms of syllable structure, phonotactic regularities, and allophonic and morphophonemic alternations" (Best and Tyler 2007:28).

is instead associated with a new L2 phonetic category. As such, discrimination of these L2 phonemes is predicted to be good, although not as good as in Two-Category/Uncategorized-Categorized assimilation. It should be noted that although the less similar L2 phoneme is expected at first to be "perceptually learned as a new L2 phonetic variant of the L1 phonological category", eventually "the learner should learn to perceive the lexical-functional contrasts between the L2 phones, and to develop a new phonological category for the phonetically 'deviant' phone" (Best and Tyler 2007:29). In addition, the more similar L2 phoneme (assimilated to the L1 phoneme on both the phonological and the phonetic level) may eventually be associated with its own phonetic category, depending on the degree of perceived similarity to the L1 phoneme.

The third possible alignment is one in which the two L2 phonemes are perceptually assimilated to the same L1 phoneme with the same goodness of fit (i.e., Single-Category assimilation in the PAM). This is the most troublesome case for the L2 learner because the L2 phonemes are assimilated to the L1 phoneme on both the phonological level and the phonetic level.[4] Given this type of assimilation, discrimination of the L2 phonemes is predicted to be poor; in fact, it is predicted that "minimally contrasting L2 words would be perceived as homophones" (Best and Tyler 2007:29).[5] The extent to which this type of L2 contrast could be learned to be distinguished depends on the goodness of fit with the L1 phoneme, but most learners are not expected to do well in this regard. Learning the contrast is predicted to require first learning a new phonetic category for one or both of the L2 phonemes; subsequently, one or two new phonological categories may be established. Although the likelihood of eventually learning this type of L2 contrast is generally low, if the "adaptive significance" of commanding the contrast is high (e.g., the contrast distinguishes many high-frequency words), then L2 learners will be put under a certain amount of "communicatively relevant pressure to perceptually learn the distinction" (Best and Tyler 2007:30), which may ultimately lead to it being learned.

[4]Presumably, this means that the two L2 phonemes are both relatively *good* exemplars of the L1 phoneme, although the PAM(-L2) states only that they are equally good or poor exemplars of the L1 phoneme.

[5]It should be noted that it is unclear, both in the PAM and in the PAM-L2, why L2 categories assimilated to an L1 category as (equally) poor exemplars of that category should always be predicted to be poorly discriminated. It is not unreasonable to think that discrimination would actually be highly affected by the configuration of the L2 categories in phonetic space relative to the L1 category (e.g., whether they are "on the same side" or "on opposite sides" of the L1 category). For example, for speakers of an L1 with a vowel system of /i, u, a, ə/, close-mid front and back vowels /e, o/ of an L2 might be perceived as equally poor exemplars of the mid central vowel /ə/; similarly, close-mid and open-mid back vowels /o, ɔ/ of a different L2 might also be perceived as equally poor exemplars of /ə/. In this case, the PAM(-L2) predicts that, at least initially, discrimination of /e/ and /o/ will be no better than discrimination of /o/ and /ɔ/. This seems improbable, however, given the difference in phonetic distance between the members of each L2 contrast. What seems more likely is that while /o/ and /ɔ/ may be discriminated poorly, /e/ and /o/ will be discriminated well, or at least better than /o/ and /ɔ/.

The fourth possible alignment is non-alignment: both L2 phonemes are significantly different from the closest L1 phoneme, such that neither is perceptually assimilated to an L1 phoneme (i.e., Uncategorized-Uncategorized assimilation in the PAM). In this case, how well the L2 phonemes are discriminated depends on both their degree of similarity to L1 phonetic categories and the degree of phonological overlap between the constellations of L1 phonetic categories to which they are respectively deemed similar (i.e., whether the L2 phonemes are similar to L1 phonetic categories associated with the same L1 phonological category or different L1 phonological categories). It is predicted that "if each of these uncategorized L2 phones has similarities to *different* sets of L1 phones, that is, they are relatively distant from one another within L1 phonological space, then the listener should easily recognize relevant L2 lexical-functional differences, and two new L2 phonological categories should be perceptually learned", resulting in good discrimination of the contrast; on the other hand, "if the uncategorized L2 phones are perceived as similar to the same set of L1 phonemes, that is, are *close* to each other in phonological space, then the listener should find them difficult to discriminate, and should not easily perceive relevant L2 lexical-functional differences" (Best and Tyler 2007:30). In the latter case, it is expected that only one new L2 phonological category will be learned. Whether the two L2 phonemes associated with this one L2 phonological category are distinguished or merged at the phonetic level then depends on their phonetic similarity (e.g., absolute overlap in the sets of L1 phones to which they bear similarities).

A fifth possible alignment is analogous to the Non-Assimilable type of assimilation in the PAM: both L2 phonemes are so different from any known speech sound that they are perceived as non-speech (e.g., Best et al. 1988). Best and Tyler (2007) observe that little is known about how L2 acquisition proceeds in this case since there are virtually no studies of L2 learning situations such as L1 English speakers learning Zulu. They speculate that "non-assimilated sounds might eventually be perceptually incorporated into the phonological space of the L2 listeners as uncategorized speech sounds, possibly resulting in the perceptual learning of 1–2 new phonological category(s)"; however, it is also possible that "L2 learners may never incorporate these sounds into their phonological space" (Best and Tyler 2007:31). In the latter case, L2 learners of Zulu, for example, might still be able to pronounce clicks, but probably as the product of non-linguistic oral gymnastics rather than as properly coarticulated consonants. In either case, the implication is that discrimination of the L2 phonemes will vary according to the magnitude of the auditory (and perhaps gestural) differences between the phonemes, but is likely to be quite good.

The PAM-L2 thus applies the principles of the PAM to L2 speech perception, making largely the same predictions regarding the perceptual consequences of different types of L2-to-L1 assimilation. However, the PAM-L2 differs significantly from both the PAM and the SLM by incorporating the phonological level of both L1 and L2 into judgments of L1-L2 similarity. Perceptual assimilation in this model, unlike perceptual assimilation in the PAM or equivalence classification in the SLM, may occur at the phonological level as well as the phonetic or gestural level. By making reference to three distinct levels of structure, the PAM-L2 assumes a multi-level model of bilingual speech relating different degrees of linguistic abstraction across languages. Such a model is implicit or explicit in much of the work on bilingual phonology (including work in the SLM) and is outlined in the section that follows.

2.3.4 Linking L1 and L2 in a Bilingual System

To account for a range of findings on L1-L2 phonological interaction in bilinguals (see Mack 2003 for an excellent review), Laeufer (1996) laid out a typology of bilingual phonological systems combining aspects of the bilingual lexical/conceptual model of Weinreich (1953) with the speech production model of Keating (1984). Based on findings for VOT production in L1 and L2 voiceless stops, this typology aims to model the variety of bilingual production patterns in terms of differences in the language specificity of a category at a given level of representation (i.e., whether the category instantiates a sound in only one language or in both languages). This representational mechanism distinguishes between bilingual phonological systems showing different degrees of cross-language influence.

Laeufer's typology assumes a tripartite, hierarchical model of speech production with a phonological (i.e., phonemic) level at the top, a phonetic (i.e., allophonic) level in the middle, and a realizational (i.e., motoric) level at the bottom. In this model, a phonological category corresponds essentially to the abstract phoneme, while a phonetic category corresponds to the environmentally conditioned allophone, "a level of representation similar to the systematic phonetic level in generative models of phonological analysis" (Laeufer 1996:327), which specifies general articulatory and acoustic properties. For example, voiceless bilabial stops in a language like English (where they are produced with long-lag VOT) and voiceless bilabial stops in a language like Spanish (where they are produced with short-lag VOT) are both represented as /p/ at the phonological level, while the macro difference in VOT—that English /p/ is aspirated, while Spanish /p/ is not—is captured at the phonetic level, where English /p/ is represented as [p^h], while Spanish /p/ is represented as [p]. Whereas the phonetic level specifies phonetic characteristics in general, such as presence versus absence of aspiration, the realizational level spells out these characteristics in detail with phonetic realization rules. These rules "provide sensorimotor phonetic detail, such as exactly how long after release voicing starts, for instance, in English as opposed to German long-lag voiceless stops (i.e. the exact amount of aspiration)" (Laeufer 1996:327–328). In this respect, the realizational level in this model might be considered an analogue of the gestural level in the PAM-L2.

Cross-language influence is represented in this model primarily by the merger of L1 and L2 categories at different levels of representation, with the extent of cross-language influence increasing as categories are merged at lower levels of representation. In addition, the language specificity of a merged category plays a role in determining what kind of system results. Three main kinds of bilingual phonological systems (coexistent systems, merged systems, and super-subordinate systems) are delineated in this way:

> The distinguishing properties of the three kinds of systems at the phonological and phonetic levels...are thus two versus one series of stops and, in the case of a single series, a mixed L1-L2 versus a simple L1-specific representation. At the phonetic realization level...the distinguishing features are (a) different versus identical L1 and L2 realizations and, in the case of identical renditions, fused versus L1-specific realizations; (b) native-like versus non-native-like L1 and L2 and, in the case of non-native-like L2, smaller versus larger deviation from the L2 phonetic norm. (Laeufer 1996:331)

Thus, coexistent, merged, and super-subordinate systems may be differentiated by the presence versus absence of merger at the phonological, phonetic, and realizational levels and, in the case of merger, by the distance of the merged L1/L2 representation from the component L1 representation (i.e., whether the merged representation is closer to L1 versus intermediate between L1 and L2). The kind of system manifested by a bilingual is influenced by several interacting extralinguistic variables, including "the amount of past and...present intensive exposure to the speech of native L2 speakers, the level of proficiency and dominance profile at the time a speaker is tested, and the age at which L2 acquisition began" (Laeufer 1996:339). Figure 2.1 schematizes the three kinds of systems with respect to voiceless stop production for an L1 with short-lag voiceless stops (e.g., [p]) and an L2 with long-lag voiceless stops (e.g., [p^{h}]).[6]

The first kind of system, schematized in Figure 2.1a, is the COEXISTENT system, in which similar L1 and L2 sounds are represented distinctly at all three levels. Coexistent systems are observed in "very proficient bilinguals who are actively using both languages" (Laeufer 1996:339), typically individuals who acquired the two languages simultaneously as L1s or who acquired L2 very early in life. These systems are associated with native-like production in one language as well as a "high level of proficiency

[6]Though based on Laeufer (1996:329), the tree schematics presented in Figure 2.1 differ from the trees of Laeufer (1996) in several ways. First, they eliminate unidirectional and bidirectional cross-linguistic arrows between L1 and L2 categories within the same level of representation (which are largely redundant with the representational mechanism of category merger), as well as language indices to a "compound", or merged, category of L1 and L2. Moreover, symbols are italicized, rather than put in parentheses, when the phonetic norm of the given language is not reached. In addition, square brackets are used to represent the phonetic level (consistent with conventions for allophonic-level transcription in the International Phonetic Alphabet), while braces are used to represent the realizational level. Finally, Laeufer (1996) sketched a number of theoretically impossible systems, which are not discussed here because they can be uniformly ruled out by positing a downward branching requirement, which prevents nodes from converging at a lower level (i.e., the relation between nodes at a higher level and those at a lower level can be one-to-one or one-to-many, but not many-to-one).

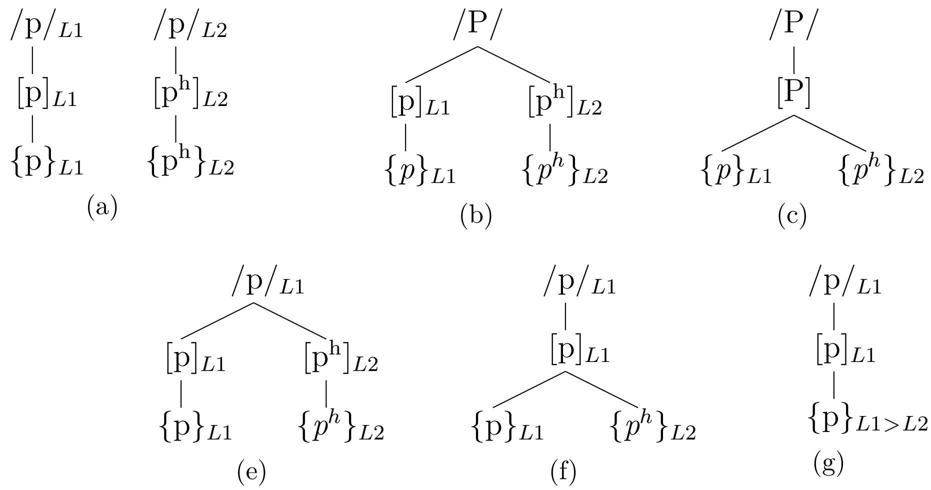


Figure 2.1: A typology of the different possible bilingual systems of voiceless stops, based on Laeufer (1996:329). Coexistent systems are schematized in 2.1a, merged systems of Type I–III in 2.1b–2.1d, and super-subordinate systems of Type I–III in 2.1e–2.1g. Slashes denote the phonological (phonemic) level; brackets, the phonetic (allophonic) level; and braces, the realizational (motoric) level. Capitals denote a mixed, non-language-specific representation, while italics indicate that the phonetic norm of the given language is not reached.

reached in the other language, with either equally native-like...or near-native-like" production (Laeufer 1996:338).[7] Bilinguals with a coexistent system of L1 and L2 voiceless stops, for example, are expected to produce stops in both languages with native-like VOT, since the separation of L1 and L2 at all three levels of representation allows for no cross-linguistic interference. This formal separation of close L1 and L2 sounds might be thought of as representing a mental separation between these sounds: the two sounds do not influence each other because they are considered totally different entities. Note that there is considerable evidence of bilinguals' production of L1 or L2 being native-like (e.g., Fokes, Bond, and Steinberg 1985; Mack 1989). However, many of these studies examined production in only one of the two languages. Partly owing to this fact, there is less evidence of bilinguals' production of both languages being native-like. Nonetheless, such production patterns have been attested, usually in early bilinguals (e.g., Williams 1977; Kang and Guion 2006).

[7] To account for cases of bilingual production in which L1 is native-like but L2 is not quite native-like (e.g., Caramazza, Yeni-Komshian, Zurif, and Carbone 1973; Flege and Eefting 1987a), Laeufer (1996) also posited a Type II coexistent system. It is unclear, however, how interference could arise within this system, which is representationally identical to the system shown in Figure 2.1a. Thus, the data she meant to account for with a Type II coexistent system are analyzed here as the product of the super-subordinate system that is conspicuously missing from her typology—a Type I super-subordinate system (Figure 2.1e).

The second kind of system is the MERGED system, in which similar L1 and L2 sounds share a representation at one or more levels. Merged systems are observed in "fairly advanced to advanced late bilinguals with L1 or L2 dominance" and in "early bilinguals with less recent or less intensive exposure to L2" (Laeufer 1996:339). Three types of merged systems are identified, which differ in terms of the lowest level at which categories are merged across languages. Merger may occur only at the phonological level (Type I: Figure 2.1b); from the phonological level down to the phonetic level (Type II: Figure 2.1c); or all the way down to the realizational level (Type III: Figure 2.1d). What merged systems have in common is "the presence of bi-directional cross-linguistic influence" (Laeufer 1996:328), which increases in strength from Type I to Type III. This cross-language interaction results in merged systems of all types being characterized by non-native-like production, in patterns that are largely consistent with the predictions of the SLM. Bilinguals with a Type I merged system produce substantial distance between similar L1 and L2 sounds; however, this distance usually arises from dissimilation between the sounds, with the result that neither is produced as native-like (e.g., Flege and Eefting 1987b). Bilinguals with a Type II merged system produce relatively little distance between similar L1 and L2 sounds due to merger at the phonetic and phonological levels (e.g., Williams 1979; Major 1992), while bilinguals with a Type III merged system produce similar L1 and L2 sounds virtually identically (with phonetic parameters intermediate between the two sounds) due to merger at all levels (e.g., Flege 1987b). The occurrence of dissimilation in a Type I merged system is consistent with the SLM's prediction of dissimilation between close L1 and L2 sounds that are each associated with a unique phonetic category. In addition, the inaccurate production of close L1 and L2 sounds in Type II and Type III merged systems is consistent with the SLM's predictions for "similar" L2 sounds that do not induce the formation of a new phonetic category, as well as for the L1 sounds to which they are judged similar.

The third kind of system is the SUPER-SUBORDINATE system, in which a (subordinate) L2 sound is parasitic on (superordinate) L1-specific representations. The L1-specificity of higher-level representations is a central feature of the super-subordinate system distinguishing it from merged systems: while merged systems have "compound (i.e. mixed) phonological representations", super-subordinate systems have "a single L1-specific phonological representation" (Laeufer 1996:330). Super-subordinate systems are observed to arise in "speakers with less exposure to native L2 speech and/or less overall proficiency", "speakers with less recent exposure to L2", and "speakers whose L2 has fossilized at an elementary level of acquisition" (Laeufer 1996:339). Laeufer posited two types of super-subordinate systems, to which a third is added here; all are characterized by L1-specific phonological categories. The types differ in terms of the L1 category on which a similar L2 sound is parasitic. The L2 sound may be associated with a unique phonetic realization that draws upon information contained within the L1-specific phonetic level (Type II: Figure 2.1f) or simply co-opt the phonetic realization of the L1 sound (Type III: Figure 2.1g). Alternatively, the

L2 sound may be associated with its own phonetic category, which does not cause dissimilation between the L2 sound and the corresponding L1 sound because of the L1-specificity of the phonological level (Type I: Figure 2.1e). What Type II and III super-subordinate systems have in common is "native-like L1, and L2 values which are very similar" (Laeufer 1996:338). In the case of Type II super-subordinate systems, close L1 and L2 sounds are distinguished with a minimal amount of differentiation, while in the case of Type III super-subordinate systems, they are produced virtually identically, both with L1-like realizations. Just as with coexistent systems, bilingual studies often provide substantial support for only one of the two main features of the super-subordinate system—either the aspect of L1-influenced L2 production (e.g., Port and Mitleb 1983; Gass 1984) or the aspect of native-like L1 production.

Laeufer's typology of bilingual phonological systems formalizes many of the important elements of the PAM(-L2) and the SLM. Merger of L1 and L2 sounds corresponds essentially to the concept of perceptual assimilation in the PAM(-L2) and to the concept of equivalence classification resulting in perceptually linked diaphones in the SLM. In terms of the PAM-L2, when there is no perceptual assimilation of an L2 sound, the L2 sound and the closest L1 sound may be said to be in a coexistent system. In terms of the SLM, when an L2 sound evades equivalence classification and is considered to be "new", it causes the establishment of a separate phonetic category, thus resulting in a coexistent system or a Type I merged or super-subordinate system (depending on whether or not a separate phonological category is also established for the L2 sound). A Type I merged or super-subordinate system may also result from perceptual assimilation of an L2 sound at the phonological level. Perceptual assimilation at the phonetic level leads to a merged or super-subordinate system of Type II, while perceptual assimilation at the gestural level leads to a merged or super-subordinate system of Type III. Whether perceptual assimilation leads to a merged system or a super-subordinate system presumably depends on the degree of similarity between the L2 sound and the L1 sound to which it is assimilated; however, this question falls outside the purview of the PAM(-L2) and is thus not addressed in the model. Like perceptual assimilation, equivalence classification at the phonetic level is consistent with a merged or super-subordinate system of Type II. However, given the SLM postulate that L1 phonetic categories continue to develop in adulthood under the influence of L2 sounds that are identified with them, the SLM seems to predict that equivalence classification eventually leads to a merged system.

In sum, assuming a three-tier model of speech production predicts at least seven different types of bilingual phonological systems, which fall into one of three kinds: coexistent, merged, and super-subordinate. In coexistent systems, similar L1 and L2 sounds are both produced as native-like. On the other hand, in merged systems, neither sound is produced as native-like, while in super-subordinate systems, only the L1 sound is produced as native-like. This framework formalizes the concept of cross-language interference between L1 and L2 sounds (cf. perceptual assimilation in the PAM, equivalence classification in the SLM) in terms of shared representation at any of the phonological, phonetic, and realizational levels and is largely compatible with the tenets and predictions of the PAM(-L2) and the SLM.

2.3.5 Summary

The most comprehensive view of bilingual production is provided not by any one of the models discussed above, but by joint consideration of the SLM and the PAM-L2. Though the SLM and the PAM differ in terms of the object of study—the SLM focusing on the L2 phonology of advanced bilinguals, and the PAM focusing on non-native perception of L2 by naive listeners—both models predict that equivalence between L1 and L2 sound categories, at least at the beginning stages of L2 acquisition, should be established based on phonetic similarity. The SLM posits that L2 categories are linked to L1 categories perceptually via the mechanism of equivalence classification, which has most often been interpreted in acoustic phonetic terms. On the other hand, the PAM claims that perceptual assimilation is based on articulatory-gestural similarity. Given that acoustic similarity and articulatory similarity are highly interrelated, both models are capable of explaining the vast majority of category equivalences between sounds that show a high degree of phonetic similarity. What both models fail to account for, however, are category equivalences that seem instead to draw upon relationships of phonological similarity. The PAM-L2 expands upon the PAM and the SLM by allowing for cross-language equivalences to be established at the phonological level. Nonetheless, the PAM-L2, like the PAM, remains a model of perception, not production. Thus, whereas the SLM is the only one of these models that makes explicit predictions regarding L2 production, the PAM-L2 is the only model that addresses category equivalence at the phonological level; in this way, they are both relevant to the study of bilingual speech production. These models assume a three-level view of speech production, in which cross-language linkage between L1 and L2 sounds may be conceptualized in terms of merger at the phonological, phonetic, and/or realizational level. In this framework, the way in which L1 and L2 sounds share a representation predicts the way in which they will be produced, both in relation to each other and with respect to native phonetic norms. Several different bilingual phonological systems are possible, most of which involve merger of L1 and L2 at one or more levels and, therefore, cross-linguistic influence. Such cross-language phonological interaction is documented in diverse situations of language contact, which are reviewed below.

2.4 Phonological Interaction in Language Contact

Although the literature on bilingual phonology provides evidence for each of the bilingual phonological systems described in Section 2.3.4, there is an imbalance of evidence in favor of merged systems and bi-directional cross-linguistic influence. As noted above, this imbalance is attributable in part to bilingual studies often being one-sided, examining performance in only one language rather than both languages. However, the recurring finding of bilingual production patterns that are significantly different from monolingual norms suggests that it may simply be more likely for bilinguals to have a merged system than a coexistent or super-subordinate system. A bias toward merged systems seems reasonable, as the conditions that appear to be required for a coexistent system to surface (early, consistent exposure to L2; current, frequent use of both L1 and L2) and the conditions that would limit speakers to a super-subordinate system (little exposure to L2, both recently and overall) are generally less likely to be achieved than conditions that are intermediately conducive to bilingualism (interrupted or late exposure to L2; less frequent use of L1 or L2).

Thus, perhaps due to the high probability of obtaining acquisition conditions favoring a merged system, phonological studies of "bilinguals" (broadly construed) have largely shown interaction between L1 and L2 in individuals who have significant proficiency in both languages. Phonological interaction between two languages has been described in a variety of linguistic situations. This section reviews findings from studies of three groups: last speakers of obsolescent languages, heritage speakers of minority languages, and bilingual speakers of non-endangered languages. These groups differ in terms of amount of exposure to and current use of L1 in comparison to L2, yet across them there is a common finding: L1 phonology can be significantly influenced by L2 phonology.

2.4.1 Sound Change in Language Obsolescence

Environments favoring L1 attrition (Seliger and Vago 1991) encourage L2 influence on L1, and there is perhaps no linguistic situation more conducive to such cross-linguistic effects than language obsolescence. In a wide-ranging survey of severely endangered languages, Campbell and Muntzel (1989) developed a typology of language obsolescence and the sorts of change processes that can occur in obsolescing languages. One of the most common types of obsolescence is "gradual death", a situation in which an endangered language (L1) is eventually lost due to increasing bilingualism in a dominant contact language (L2), which eventually comes to be used in all communicative contexts. In such an obsolescing L1, there are three main patterns evident in the types of phonological changes that occur (Andersen 1982:95). First, fewer phonological distinctions are made overall than at more viable stages of the L1. Second, phonological distinctions common to the obsolescing L1 and the dominant L2 are preserved. Finally, phonological distinctions with a high functional load are maintained longer than those with a low functional load. Of these three patterns, both the first and second ones have the potential to make L1 more similar to L2. The first pattern, a loss of structure, is most often convergent with L2, since the structure lost is usually one not found in the L2; through the loss of structure particular to L1, the obsolescing L1 becomes more similar to the contact L2. On the other hand, in reinforcing the structure that L1 and L2 have in common, the second pattern may result in L1 converging with L2 by causing shared elements to be kept at the expense of unshared elements.

Citing much of Campbell's previous work in this area, Campbell and Muntzel (1989:186–187) identified many cases of convergent phonological change. One example occurs in the language Pipil (Southern Uto-Aztecan, Aztecan), whose speakers have for the most part neutralized a vowel length contrast not found in the dominant language, Spanish, leaving just short vowels. In Chiltiupan Pipil, the alveolar affricate /ts/ has furthermore merged with the fricative /s/. In the case of Tuxtla Chico Mam (Mayan, Mamean), a contrast between velar and post-velar/uvular plosives, again not found in dominant Spanish, has disappeared, leaving just velars. In addition, Finnish Americans have been shown not to faithfully produce the vowel length contrast, singleton-geminate consonant length contrast, or front rounded vowels of European Finnish—all phonological features that are absent from the dominant language, English. Goodfellow and Alfred (2002:215) and Goodfellow (2005:134–138) documented several other examples of convergent phonological change in younger generations of Kwak'wala (Wakashan, Northern) speakers. These speakers have lost several classes of Kwak'wala sounds that are absent from English, either omitting them or replacing them with more familiar sounds from English: glottalized consonants are replaced by plain pulmonic consonants, uvulars are replaced by velars, velar fricatives are omitted, and lateral affricates are replaced by /gl/ clusters. Bullock and Gerfen (2004a,b, 2005) reported yet more examples of convergent change in the endangered language variety of Frenchville French, although not all oppositions potentially at risk of merger were found to merge (Bullock, Dalola, and Gerfen 2006).

Campbell and Muntzel (1989) described these sorts of externally motivated (i.e., L2-influenced) changes in L1 as predictable or expected. What they have in common is the loss of structures in the obsolescing language that are not present in the dominant language. Campbell and Muntzel also enumerated several other categories of phonological change that they described as "of uncertain predictability", which include the overgeneralization of unmarked features and the overgeneralization of marked features. The overgeneralization of unmarked features can result in the types of convergent change cited above for Pipil, Tuxtla Chico Mam, American Finnish, and Kwak'wala (short segments are indeed less marked than long segments, velars less marked than uvulars, schwa and back rounded vowels less marked than front rounded vowels, and pulmonic consonants less marked than glottalized consonants). The internal effect of unmarkedness/naturalness and the external effect of a dominant language on the loss of structure are therefore indistinguishable when the structure lost is a marked structure present in the obsolescing language and absent from the dominant language; either or both of these effects may be responsible for the apparently convergent change.

On the other hand, the overgeneralization of marked features normally results in divergent change. Campbell and Muntzel (1989) cited Jumaytepeque Xinca (isolate) as one case of a marked form being overgeneralized, with the result that the language has diverged from the dominant language that does not have the marked form. In this case, Jumaytepeque Xinca has a rule glottalizing consonants in specific environments, but some speakers have lost this rule and generalized the relatively marked glottalized consonants to all environments over the relatively unmarked plain consonants. Teotepeque Pipil provides another example of this kind of overgeneralization. In this case, voiceless [l̥] used to be a word-final allophone of voiced [l], but speakers came to generalize this relatively marked segment to all environments. Although Campbell and Muntzel (1989:189) reasoned that such divergent changes "are internal to the structure of the obsolescent language in that they appear to have no direct analog in the dominant language", Woolard (1989) countered that these sorts of divergent changes may actually be externally motivated: a marked structure not present in the dominant language is exaggerated in the obsolescing language to differentiate it from the dominant language. In this way, the divergent change may serve as a symbolic act of distancing from the dominant language by speakers who want to "emphasize their differentness from the dominant group" (Thomason 2001:230), a motivation reminiscent of the sociolinguistic situation in Martha's Vineyard described by Labov (1963). Here, enhancement of the local English dialect feature of vowel centralization in the speech of young people correlated with how strongly they identified as residents of Martha's Vineyard: the more strongly they identified as locals, the greater the degree of vowel centralization in their speech.

Thus, while it is possible for divergent change to occur in cases of language obsolescence, convergent change remains the more commonly attested type of change, often resulting in the merger of two phonological categories that do not contrast in the dominant language. Given that these sorts of mergers are widely attested, it should come as no surprise that they do not form a homogeneous class. One feature used to differentiate them is the path to merger, which may occur either via "transfer" or "approximation" (Trudgill and Foxcroft 1978). In the case of transfer, two phonemes merge via the first phoneme categorically changing to the second phoneme in more and more words containing the former phoneme (cf. super-subordinate systems in Laeufer 1996); in this case, the merger is accomplished by the unidirectional transfer of one phoneme to another in a process that "involves...a form of lexical diffusion" (Trudgill and Foxcroft 1978:73), which is "not consistent with a result that shows an intermediate phonetic form" (Labov 1994:324). In the case of approximation, however, two phonemes merge as their individual phonetic spaces approach (i.e., approximate) each other; here both phonemes typically shift, resulting in a merged category with a phonetic space intermediate between the original phonemes (cf. merged systems in Laeufer 1996). According to Labov, approximation may also result in a merged phoneme with approximately the same phonetic space as one of the original phonemes; similar to transfer, then, the final result in this sort of approximation is not an intermediate phonetic form. In addition to these two merger types, Labov (1994:321–323) added a third type, expansion, in which the phonetic space of the merged category, rather than being intermediate between the original categories or coincident with one of them, spans the phonetic spaces of both.

Campbell and Muntzel (1989) reviewed three other types of sound change that can occur in obsolescing languages: development of variability in rules, under- and overgeneralization of rules, and replacement of stigmatized forms with prestige forms. First, variability may develop in the application of phonological rules. Rules that used to be obligatory may apply optionally, show substitutions, or simply be lost. The case of optional rule application usually results in a situation of free variation between forms that have resulted from the rule and those that have escaped it. For example, consonant gradation rules in standard Finnish that voice stop consonants in certain environments are not applied consistently in American Finnish, producing free variation between voiced and voiceless stops in environments where only voiced stops would occur in standard Finnish. Second, phonological rules may be undergeneralized on the one hand and overgeneralized on the other. In the case of Teotepeque Pipil mentioned above, a rule which devoiced sonorants word-finally has been overgeneralized for /l/, resulting in voiceless [l̥] in all environments, but undergeneralized for /w, j/, resulting in voiced [w, j] in all environments. Finally, as noted by Labov (1994:321), foreign phonemes from the dominant language may replace native phonemes in an obsolescing language when the foreign phoneme is more prestigious (and especially when the native phoneme is stigmatized). For instance, the unnatural sound change of /s̥/ > /r/ in Teotepeque Pipil is most likely attributable to the fact that the former is a stigmatized form in the regional Spanish, whereas the latter is a prestige form. This shift can thus be described as an externally motivated change.

Three recent studies have applied the typology of change laid out above to current situations of language endangerment. In an acoustic and articulatory study of Northern Paiute (Uto-Aztecan, Western Numic), Babel (2009a) documented two kinds of sound change in the language. First, while a three-way laryngeal contrast has been maintained in each of three generations of speakers, the phonetic realization of this contrast differs across generations, with increased subphonemic variation in the youngest generation. Second, the place of articulation of the language's sibilant has shifted from a palatalized post-alveolar to a plain alveolar (i.e., English /s/), while a more palatalized allophone has been replaced by the English palato-alveolar /ʃ/ in the youngest generation. Based upon these results, Babel hypothesized that contrasts based on timing relationships (e.g., laryngeal contrasts) were more likely to undergo sound change via approximation, while more categorical contrasts (e.g., consonantal place contrasts) were more likely to undergo sound change via transfer.

However, counterevidence to Babel's hypothesis was found by Chang (2009a), who observed that consonantal place contrasts in Southeastern Pomo (Northern Hokan, Pomoan) seem to be undergoing change via approximation. In a cross-generational investigation of speech production by the last Southeastern Pomo speakers, Chang argued that convergent and divergent sound change can co-occur within an individual speaker under the same external influences from a contact language, such that the obsolescing language becomes simultaneously more similar to and more different from the contact language. The results of a series of acoustic case studies showed that while non-English place contrasts (velar/post-velar, dental/alveolar) are still produced in the youngest generation of fluent speakers, the amount of differentiation produced between the two members of each contrast has decreased—a convergent change suggestive of external influence from English. Similarly, a divergent change eliminating rhotics wholesale from the consonant inventory seemed also to have an external motivation—in this case, a hypercorrective reaction to external influence from English.

Haynes (2010) identified yet another possible mechanism of sound change in endangered languages—namely, "areal hypercorrection", whereby speakers may adopt salient phonological features of other languages in the geographic vicinity (which are typically also endangered). This study compared production of Numu (Oregon Northern Paiute) in an imitation task by a group of fluent Numu speakers to production by two groups of adult non-speakers: English speakers from within the community where Numu is spoken (who had either ambient or direct exposure to Numu) and English speakers from outside this community (who had no significant prior exposure to Numu). The results of this comparative analysis suggested that "insider" participants from the speech community had a general advantage in production over "outsider" participants, and that this advantage was mediated by previous experience, with insiders who had direct exposure to the language often outperforming insiders who had only ambient exposure. In a series of acoustic case studies, evidence was found of three previously proposed routes of change in endangered languages: interference from a dominant language (transfer), incorporation of universal language features (regularization), and enhancement of socially salient language features (hypercorrection). Significantly, while outsiders' production showed only transfer or regularization, insiders' production also showed evidence of hypercorrection, consistent with the differences in sociocultural ties to the language between the two groups. Some insiders, moreover, were found to produce segments that were neither English nor Numu, but that were characteristic of other local Native American languages. This result led Haynes to propose the change type of areal hypercorrection, which, along with regular hypercorrection, was found in an accent rating task to be especially detrimental to perception by fluent speakers.

Taken together, the findings of work on sound change in endangered languages suggest that an L1 is susceptible to profound influence from an L2 when its usage is severely diminished due to social factors. Many of the studies discussed above describe cases of bilinguals for whom the endangered language is indeed L1 (the first language to which they were exposed), yet they repeatedly find that the L1 speech of these individuals shows significant changes at both a phonemic and subphonemic level with respect to the speech of earlier generations, who had socially more robust experience with the language. Often these changes are due to external influence from a dominant L2, which came to be used more often than the L1 as the number of fluent

L1 speakers rapidly decreased. The direction of such L2-influenced change is usually convergent with the L2, though it is also possible for the external influence to cause divergent change accentuating socially salient differences between the two languages. L2 influence on L1 is unsurprising in the situation of language endangerment, which is particularly conducive to phonological attrition. However, similar results have also been reported for individuals with non-endangered L1s, including heritage speakers of minority languages.

2.4.2 Intermediate Phonology in Heritage Speakers

Similar to the case of language endangerment, the ecological context of an L1 acquired as a heritage language (HL) differs in many ways from that of an L1 acquired in the canonical way (i.e., early, consistent exposure from birth, followed by frequent, continuous use throughout life in a community of many other L1 speakers). Such differences in context might be expected to lead to differences in command and usage of the language, but there are comparatively few technical studies that focus on HL speakers as a group. Although there exists a large body of scholarship on the linguistic competence of child L1 and adult L2 acquirers, researchers have only begun to examine the linguistic knowledge of HL speakers—that is, individuals whose current primary language differs from the language they spoke or only heard as a child (the HL). As fully comparable neither to native speakers (who acquired the language completely as children) nor to late L2 learners (who acquired the language as adults, typically incompletely), HL speakers form somewhat of an intermediate group, a group of interest because they have an often rich knowledge of their HL, even when they do not actively speak the language. Typical HL re-learners are predicted to have acquired "nearly 90% of the phonological system" and "80% to 90% of the grammatical rules" of the HL—a significantly more extensive command of the language than second-year college L2 learners (Campbell and Rosenthal 2000:167). Indeed, childhood experience with a minority language, even if merely overhearing, has been found to provide a significant boost to a speaker's phonological production and perception of that language later in life in comparison to L2 learners with no prior experience (Tees and Werker 1984; Knightly, Jun, Oh, and Au 2003; Oh, Jun, Knightly, and Au 2003). HL speakers also tend to be more native-like than L2 learners in their morphosyntax, although they nonetheless pattern differently from native speakers (Montrul 2008; Au, Oh, Knightly, Jun, and Romo 2008; Polinsky 2008).

Studies of HL phonology have been conducted on a number of languages, including Armenian (Godson 2003, 2004), Korean (Au and Romo 1997; Oh, Au, and Jun 2002; Oh et al. 2003; Au and Oh 2009), Russian (Andrews 1999), and Spanish (Au and Romo 1997; Au, Knightly, Jun, and Oh 2002; Knightly et al. 2003; Oh and Au 2005; Au et al. 2008), the majority of this research coming out of joint work by Au, Jun, Knightly, Oh, and Romo on HL speakers of Korean and Spanish. In their series of studies, which included acoustic measures such as VOT and degree of lenition, holistic measures such as overall accent ratings, and perceptual measures such as phoneme identification accuracy, the recurring theme is that HL speakers tend to have a phonological advantage over L2 learners. However, whether HL speakers show an advantage just in perception or in both perception and production of the HL seems to be related to the nature of their HL experience. In this regard, Au and colleagues distinguished between "childhood hearers" and "childhood speakers". Knightly et al. (2003), for example, focused on childhood overhearers of Spanish—Spanish speakers who had regular childhood experience with overhearing Spanish, but not with speaking or being spoken to—and found that these childhood overhearers were measurably better than L2 learners at producing individual Spanish phonemes as well as whole Spanish narratives. Similarly, Oh et al. (2003) found that HL speakers of Korean had a phonological advantage over L2 learners; however, they examined not only childhood hearers, but also childhood speakers who spoke Korean regularly during childhood. Comparing these two HL groups, they found that while childhood speakers were measurably better than L2 learners in both perception and production of Korean, childhood hearers were better than L2 learners only in perception. This discrepancy with the results of Knightly et al. (2003) was attributed to two possible factors: the difference in average duration of HL re-learning (longer in the case of the HL Spanish speakers) and the difference in complexity between the two contrasts examined (a two-way laryngeal contrast in Spanish between voiced and voiceless stops vs. a three-way laryngeal contrast in Korean among lenis, fortis, and aspirated stops/affricates). In short, the findings of Au and colleagues suggested that previous HL speaking experience confers an advantage in both production and perception of the HL, and that previous HL listening experience confers an advantage in perception of the HL.[8] However, the benefit conferred by HL listening experience in production of the HL appears to be mediated by additional factors.

Although studies of HL phonology have investigated the relative authenticity of HL speakers' production as compared to late learners', few have explicitly examined the question of categorical merger (i.e., whether HL speakers merge different sound categories rather than producing them distinctly). Suggestive results were obtained by Godson (2003, 2004), who found that HL speakers of Western Armenian showed some influence of close English vowels in their pronunciation of the Armenian back vowels, but not necessarily to the point of merger. The issue of merger was also explored in work on HL speakers of Mandarin Chinese by Chang, Haynes, Yao, and Rhodes (2009, 2010). Given the SLM's hypothesis that new phonetic categories are more likely to be formed for L2 sounds at younger ages of learning, these studies tested the hypothesis that HL speakers, due to their childhood experience with two

[8]Note that contradictory null results have emerged from the work of Pallier and colleagues (Pallier, Dehaene, Poline, LeBihan, Argenti, Dupoux, and Mehler 2003; Ventureyra and Pallier 2004; Ventureyra, Pallier, and Yoo 2004), who, examining subjects from a different HL situation, have failed to find a perceptual, or even low-level neural, advantage for the HL in individuals with early HL exposure (Korean adoptees in France).

languages, would outperform late L2 learners in effecting contrast: language-internal phonological contrast between phonemic categories (e.g., Mandarin /ʂ/ and /ɕ/), as well as cross-linguistic phonetic contrast between similar, yet acoustically distinct categories of different languages (e.g., Mandarin /u/ and English /u/). To this end, production of Mandarin and English by Chinese American HL speakers of Mandarin was compared to that of native Mandarin-speaking late learners of English and native American English-speaking late learners of Mandarin in three experiments. In Experiment 1, back rounded vowels in Mandarin and English were produced distinctly by all groups, but the greatest separation between similar vowels was achieved by HL speakers. In Experiment 2, Mandarin aspirated plosives and English voiceless plosives were produced distinctly by native Mandarin speakers and HL speakers, who both put more distance between them than late learners of Mandarin. In Experiment 3, the Mandarin retroflex and English palato-alveolar fricatives were distinguished by more HL speakers and late learners than native Mandarin speakers. Thus, overall the hypothesis was supported: across experiments, HL speakers were found to be the most successful at simultaneously maintaining language-internal and cross-linguistic contrasts, a result that was attributed in large part to a close approximation of phonetic norms that occurs during early exposure to both languages.

It should be noted that while work on HL phonology has generally been concerned with the ways in which HL speakers differ from L2 learners, the data available in previous studies also allow for comparisons to native speakers. These data suggest that, consistent with the intermediate nature of their linguistic experience, HL speakers tend to pattern in between L2 learners and native speakers in production of the HL.[9] As discussed above, they have often been found to be significantly more native-like than L2 learners; however, at the same time they are usually significantly less native-like than native speakers. For example, HL speakers of Spanish, despite native-like production of VOT, have been found to apply intervocalic lenition processes significantly less frequently than native Spanish speakers and to be rated in terms of overall accent as significantly less native-like than native speakers (Au et al. 2002; Knightly et al. 2003). HL speakers of Korean have also been found to sound reliably less native-like than native speakers, and this disparity in overall accent is reflected in some slight, but significant differences in measured VOT (Oh et al. 2002, 2003). In a similar way, though HL speakers of Mandarin have been found to pro-

[9]As of yet, examples of HL speakers patterning in between native speakers and L2 learners in production of the HL (which is usually L1) have not been paralleled in the literature by examples of such patterning in production of the dominant language (which is usually L2). Here there is less evidence, as most HL studies are concerned with production of the HL, rather than of the dominant language. However, the existing data suggest that there is less of a tendency toward intermediate patterning in the dominant language than in the HL. For example, Au et al. (2002:241) observed in a footnote that both HL speakers of Spanish and L2 learners of Spanish produced English word-initial bilabial stops with native-like VOT. Likewise, Chang et al. (2009, 2010) found no consistent differences between the English stops, fricatives, and vowels of HL speakers of Mandarin and those of L2 learners of Mandarin.

duce Mandarin /u/ much closer to native phonetic norms than L2 learners do, they still often show F_2 values that are higher than those of native Mandarin speakers (Chang et al. 2010). Thus, while HL speakers may achieve native-like production in certain aspects of the HL phonology, there is a strong tendency for them to pattern differently from native speakers in other aspects and at the level of overall accent.

In short, research on HL phonology has shown that while childhood exposure to a language generally leads to more native-like production than late exposure, the nature of HL experience, which tends to be brief, interrupted, and/or restricted, usually results in some degree of L2 influence on HL production. In this respect, speaking an HL is similar to speaking an obsolescent language; in fact, examples of HL acquisition have sometimes been included in discussions of language obsolescence (e.g., Finnish as learned by Finnish Americans; see Campbell and Muntzel 1989). What heritage languages and obsolescent languages have in common is a pattern of usage that is non-canonical for an L1: in neither case is the language spoken in all communicative contexts. An HL might be spoken only at home or to older relatives, while an obsolescent language, known only by a handful of people in the community, might hardly be spoken at all. Thus, the sociolinguistic environments of these two types of languages are similar in that their usage is significantly limited. It follows that cross-linguistic interference from a more dominant language might be especially apt to occur in these situations. However, L2 influence on L1 is also found in cases of more balanced bilingualism, in which L1 has been acquired completely and does not fall entirely out of use. The next section reviews such cases, which demonstrate that severely limited usage of L1 is not required for L1 to drift in the direction of L2.

2.4.3 Phonetic Drift in Bilinguals

The basic premise that bilinguals may show performance in each of their languages that differs significantly from monolinguals' has been supported by a number of bilingual studies. With regard to perception, Caramazza et al. (1973) found that early Canadian French-English bilinguals asked to identify stimuli in a synthetic CV continuum displayed perceptual crossovers intermediate between the crossover points of Canadian French and English monolinguals. Consistent with this result, Mack (1989) found that English-dominant, early English-French bilinguals (most of whom had learned English as L1, and all of whom were judged to sound like native English speakers) differed significantly from monolingual English speakers in identification of synthetic /d/-/t/ and /i/-/ɪ/ continua. In another well-known study, Sebastián-Gallés and Soto-Faraco (1999) found perceptual differences between two groups of early Spanish-Catalan bilinguals that differed in language dominance. They were tested on their processing of four Catalan contrasts not found in Spanish (/ɛ/-/e/, /ɔ/-/o/, /ʃ/-/ʒ/, /s/-/z/) in a forward gating task, which showed the Catalan-dominant bilinguals to have a significant advantage over the Spanish-dominant bilinguals on all contrasts except /s/-/z/. Since all these bilinguals were exposed to L2 from a very young age (∼3–4 years) and the Spanish-dominant ones that were analyzed were the most proficient in their Catalan processing, these results were taken as evidence that “there are severe limitations on the malleability of the initially acquired L1 phonemic categories, even under conditions of early and extensive exposure” and that “first language exposure modifies the speech perception system in such a way that even relatively early, intensive exposure to a new language is not sufficient to overcome the influence of L1 phonemic categories in the formation of new, non-native categories” (Sebastián-Gallés and Soto-Faraco 1999:119). Thus, the findings of perception studies have suggested that it is rare, or perhaps impossible, for bilinguals’ performance in at least one language not to differ from monolinguals’ in some way.

Differences between bilinguals and monolinguals are commonly found in the domain of production, including L1 production. When engaged in tasks encouraging both languages to be activated, even bilinguals who normally keep their languages separate have been shown to evince phonetic interaction between L1 and L2 (Bullock, Toribio, Davis, and Botero 2004; Toribio, Bullock, Botero, and Davis 2005); however, L2 influence on L1 has also been documented in bilingual studies examining one language at a time. At the level of overall accent, early L1 Korean-L2 English bilinguals in the U.S. were found by Yeni-Komshian, Flege, and Liu (2000) to produce both languages with a detectable accent. This study examined L1 and L2 production in bilinguals who differed in age of arrival (AOA) to the U.S. (range of 1–23 years) via accent ratings by native listeners. These rating data indicated that while the L2 production of participants with the earliest AOAs was significantly closer to native than that of participants with later AOAs, it was still not quite native-like. At the same time, the L1 production of participants with the earliest AOAs was also rated as accented (in contrast to participants with later AOAs, whose ratings were no different from L1 monolinguals’). Comparisons of accent ratings in L1 and L2 showed that participants with AOAs of less than 9 years had more native-like English than Korean, while participants with AOAs of greater than 12 years had more native-like Korean than English. Since even participants with the earliest AOAs had accented production, these results were argued to show that “deviations from native pronunciation result from interactions between the languages of bilinguals”, rather than from a critical period for language acquisition. Jiang (2008, 2010) found similar evidence of accentedness in the speech production of L1 Mandarin-L2 English bilinguals with AOAs to Canada in the range of 9–13 years. Two groups of bilinguals were examined: a high-use group who used L1 65% of the time on average, and a low-use group who used L1 30% of the time on average. All were young adults (18–25 years old) who had been living in Canada for 6–14 years. Vowels produced by these bilinguals were evaluated by monolingual Mandarin speakers in an accent rating task as in Yeni-Komshian et al. (2000). The rating data showed that bilinguals produced Mandarin /y/, a vowel absent from English, significantly worse than monolingual Mandarin speakers. Moreover, some bilinguals had accented productions of several other Mandarin vowels, irrespective of whether they came from the low-use or high-use groups.

Other studies have produced acoustic evidence of L1 phonetic drift as a consequence of L2 experience. In a seminal study examining cross-linguistic influence in adult L2 learners, Flege (1987b) provided evidence that the phonetic space of L1 categories changes when speakers are immersed in an L2 environment for an extended period of time. This study examined the speech production of L1 French-L2 English speakers and L1 English-L2 French speakers—both groups highly experienced in their L2 after having lived in an L2 environment for a number of years—in three case studies focusing on the realization of /t/, /u/, and /y/. Acoustic measurements in every case study showed deviance from native phonetic norms for both groups in both languages—that is, L1-influenced L2 production and L2-influenced L1 production. With regard to the voiceless stops, French /t/ (produced with short-lag VOT in native French) was produced with VOTs that were longer than native by both groups, while English /t/ (produced with long-lag VOT in native English) was produced with VOTs that were shorter than native by both groups. With regard to the back vowels, French /u/ (produced with a low F_2 in native French) was produced with F_2 values that were higher than native by both groups; conversely, English /u/ (produced with a relatively high F_2 in native English) was produced with F_2 values that were lower than native by the L1 French speakers (though not by the L1 English speakers). In contrast, French /y/, the only phoneme investigated with no phonological counterpart in English, was produced in a native-like fashion by both groups. These results were argued to follow from the classification of /t/ and /u/ by both groups as "similar" L2 sounds having clear L1 counterparts and from the classification of /y/ by L1 English speakers as a "new" L2 sound having no L1 counterpart. In terms of the SLM, the perceptual linkage of "similar" L2 sounds to L1 counterparts allowed for L2-to-L1 phonetic influence, while the "new" L2 sound was not linked to an L1 sound, thus allowing it to escape cross-linguistic influence from L1.

The finding of L1 phonetic drift in VOT has been reproduced by other researchers working with diverse bilingual populations. Major (1992, 1996) found similar evidence of L1 phonetic drift in VOT in his studies of L1 English-L2 Portuguese speakers, although he concluded that "the influence of L2 is most prevalent in casual styles of L1 and may or may not be present in formal varieties" (Major 1992:204). This research investigated the L1 and L2 production of five adult female native speakers of American English, permanent residents of Brazil who had immigrated to the country after the age of 22 years and been living there for at least 12 years. These individuals were ideal participants because they had strong motivations both for learning Portuguese and for retaining their English. They participated in three speaking tasks designed to elicit formal and casual varieties of speech: word-list reading, spontaneous sentence construction, and informal conversation. The first two tasks were completed in both English and Portuguese, while the third task was completed just in English. Acoustic measurements of VOT in voiceless stops produced by the bilingual participants indicated bidirectional cross-linguistic influence: the average VOT of their Portuguese stops was higher than found in monolingual native Portuguese speakers, while the average VOT of their English stops was lower than found in monolingual native American English speakers. Moreover, the degree of L2 influence on their L1 production was related to L2 proficiency level, an effect that was mediated by style: lower (i.e., less native-like) English VOT was significantly correlated with lower (i.e., more native-like) Portuguese VOT in casual speech, but not in formal speech. Several facts suggested that this drift in L1 was due to L2 influence, rather than other possible causes not related to L2 specifically. First, the bilingual participants continued to use L1 on a daily basis and also had strong personal and professional reasons for maintaining their L1, thus ruling out lack of L1 use as a factor. Second, the direction of the change did not show random variation, but consistent convergence with phonetic norms for L2. Finally, the bilingual participants showed no more variability in VOT than either monolingual group, making it unlikely that they were simply in a state of phonetic flux (as is often found with L1 acquirers).

L2 influence on L1 VOT was also documented by Harada (2003) in early L1 Japanese-L2 English bilinguals. These individuals produced English voiceless stops with native-like VOTs, but Japanese voiceless stops with longer VOTs than Japanese monolinguals, an effect argued to arise from dissimilation from English voiced stops for the sake of maintaining cross-linguistic contrast in a shared phonetic space. However, contradictory results were found for early L1 Korean-L2 English bilinguals by Kang and Guion (2006). In this study, bilinguals' stop production in Korean and English was analyzed in terms of VOT, f_0, and $H_1 - H_2$. Late bilinguals were found to produce English voiced and voiceless stops differently from monolingual English speakers on all three of these measures: they produced voiceless stops similarly to Korean aspirated stops and voiced stops similarly to both Korean fortis and lenis stops. Furthermore, late bilinguals produced Korean fortis stops differently from monolingual Korean speakers in terms of VOT. In contrast, early bilinguals were not found to differ from monolinguals in production of either language, suggesting that the accentedness of early Korean-English bilinguals' speech reported by Yeni-Komshian et al. (2000) likely arose from other features of their speech besides stop production.

Though the patterning of the early bilinguals examined by Harada (2003) is not consistent with that of the early bilinguals examined by Kang and Guion (2006), Harada's finding of dissimilatory drift in VOT is consistent with several previous reports in the literature. For example, Mack (1990) found evidence of VOT dissimilation in a study of an early French-English bilingual child who was exposed to both languages from infancy. Although this child's production of French voiced stops as voiceless was assimilatory to his realization of English voiced stops (which showed native-like VOT), his production of voiceless stops showed multiple stages of VOT dissimilation. On the one hand, he produced French voiceless stops with VOTs that were longer than native, in apparent dissimilation from his French voiced stops, which were produced with short-lag VOTs in the range of the native VOT norms for French voiceless stops. The now-lengthened VOT of his French voiceless stops, however, approximated the native VOT range for English voiceless stops, which thus lengthened in VOT to values far longer than native. The end result was production of both French and English voiceless stops with VOTs longer than native norms. Similar findings have been reported for L1 Japanese child learners of English by Yusa, Nasukawa, Koizumi, Kim, Kimura, and Emura (2010). In this study, Japanese children who had immersion or otherwise regular exposure to English in the context of kindergarten classes produced Japanese voiceless stops with significantly shorter VOT values than monolingual Japanese speakers, in apparent dissimilation from the long-lag VOT of English voiceless stops. The occurrence of such dissimilatory drift in individuals with very little L2 experience was taken as evidence that "L2 affects the phonetic production of L1 even when users are not proficient" (Yusa et al. 2010:583).

Flege and Eefting (1987a,b) also found dissimilatory drift of VOT in two different studies examining L1 Spanish-L2 English and L1 Dutch-L2 English bilinguals. However, as in the case of the Japanese children in Yusa et al. (2010), the dissimilation in both of these cases went in the opposite direction of the dissimilation claimed for Japanese adults by Harada (2003). Whereas Harada found that Japanese (short-lag) voiceless stops were produced with lengthened VOT, Spanish and Dutch voiceless stops, which are also characterized by short-lag VOT, were produced with shortened VOT in comparison to VOT norms of age-matched monolingual controls. This pattern held true of the child bilinguals, the early adult bilinguals, and the late adult bilinguals tested by Flege and Eefting (1987b), as well as of the bilinguals with the greatest L2 proficiency tested by Flege and Eefting (1987a). What these findings have in common is an explanation in terms of "polarization" or dispersion (Keating 1984; Laeufer 1996): an L1 sound shifts in order to maximize perceptual distance from another category in the system of contrasts. Thus, in the Japanese case, the VOT of L1 voiceless stops lengthens to dissimilate from the short-lag VOT of L2 English voiced stops, while in the Spanish and Dutch cases, the VOT of L1 voiceless stops shortens to dissimilate from the long-lag VOT of L2 English voiceless stops. It is unclear why these cases of drift, in spite of similar systems of L1 laryngeal contrast and ostensibly the same dissimilatory motivation, go in different directions. One possibility is that the VOT lengthening reported by Harada (2003) should actually be interpreted as assimilatory to English voiceless stops, similar to the example of VOT lengthening in French /t/ (Flege 1987b).

While the above studies generally examined cross-sectional samples of several speakers at one time point, Sancier and Fowler (1997) conducted a longitudinal study following one L1 Portuguese-L2 English speaker over time as she traveled between the U.S. and her native Brazil. Concentrating on this speaker's production of voiceless stops, they found that she produced shorter VOTs in both Portuguese and English immediately following months of immersion in Portuguese and, conversely, longer VOTs in both languages following months of immersion in English, although the magnitude of the difference between the two conditions was very small (on the order of 5 ms). This difference, while small, was statistically significant and, moreover, perceptible to native Brazilian listeners (though not to U.S. listeners). The phonetic drift of both languages in the direction of the ambient language was accounted for in terms of "[humans'] disposition to imitate, phonological correspondence, and preeminence of recency" (Sancier and Fowler 1997:432). First, that the ambient language has an effect on any aspect of production is due to humans' tendency to imitate what they hear. Second, that the ambient language has an effect on production in another language is due to a connection between phonologically corresponding categories in the two languages (cf. merged systems in Laeufer 1996); in other words, hearing Portuguese /t/ can affect the production of English /t/ because on the phonological level they are the same thing: voiceless coronal plosives. Finally, that L2 categories can have an effect on L1 categories even when they are acquired relatively late (as was the case with the speaker in Sancier and Fowler's study, who was already a teenager when she began learning English in earnest) is due to the heavy weighting of recently experienced exemplars in memory. In this way, recent L2 experience can affect L1 representations even though an individual may have much more cumulative experience with L1.

L1 consonants may drift spectrally as well as temporally, as shown by Peng (1993) in a study of L1 Taiwanese Amoy-L2 Mandarin speakers' fricative production. In this study, spectral analyses of frequency range and energy were conducted to compare the production of Mandarin /f/ (a "new" L2 sound in terms of the SLM) to the production of Mandarin /x/ (a "similar" L2 sound close to Taiwanese Amoy /h/) for speakers varying in L2 proficiency. It was found that although speakers showed an overall tendency to interpret the L2 Mandarin /f/ and /x/ in terms of L1 sounds ([h^w] and [h], respectively), they showed different patterns of cross-linguistic interference according to L2 proficiency level. While speakers with the lowest L2 proficiency simply substituted L1 sounds for the L2 sounds (cf. Type III super-subordinate systems in Laeufer 1996), speakers with intermediate L2 proficiency were able to produce the "new" sound /f/, but not the "similar" sound /x/, relatively accurately. Speakers with the highest L2 proficiency performed the best, producing both /f/ and /x/ relatively accurately; however, they also showed evidence of reverse interference of L2 /x/ on their L1 production.

L1 phonetic drift also occurs in the production of vowels, as seen in Flege's (1987b) findings of convergent drift in French and English /u/. Evidence of a different kind of vocalic drift was found in an acoustic study of vowel production in L1 Quichua-L2 Spanish bilinguals by Guion (2003). Four bilingual groups differing in age of L2 acquisition (simultaneous bilinguals, early bilinguals, mid bilinguals, and late bilinguals) were tested on their production of Quichua's three-vowel inventory of /ɪ, a, ʊ/ and Spanish's five-vowel inventory of /i, e, a, o, u/. While all of the simultaneous bilinguals, most of the early bilinguals, and half of the mid bilinguals were found to produce L1 and L2 vowels distinctly, most of the late bilinguals and the other half of the mid bilinguals produced an L1-like vowel space in both languages, failing to distinguish several L2 vowels from each other or from L1 vowels. Moreover, there were significant differences between the L1 vowels produced by bilinguals who had acquired the L2 vowel system accurately and the L1 vowels of a (near-)monolingual L1 speaker: L1 vowels were consistently produced by the bilinguals higher than the monolingual norms. A number of explanations for this result were considered. An interpretation in terms of assimilation to nearby L2 vowels (the L2 high vowels being higher than the L1 high vowels) was discounted because it was inconsistent with the raising of Quichua /a/, which actually took it farther away from Spanish /a/, the closest L2 vowel. An interpretation in terms of global maximization of vowel dispersion was also discarded since it was at odds with the realization of Spanish /a/, which was not produced by bilinguals as low (for the sake of differentiation from Quichua /a/), but higher than monolingual Spanish norms. In the end, the systemic raising of the L1 vowels was attributed to enhancement of perceptual distance between the L1 high vowels and the L2 non-high vowels, consistent with predictions of Adaptive Dispersion Theory (Liljencrants and Lindblom 1972; Lindblom 1986, 1990; Lindblom and Maddieson 1988; Lindblom and Engstrand 1989), which posits that phonetic systems organize to maximize the perceptual distinctiveness of each of its members while minimizing the cost to the production system (i.e., articulatory effort). This account explains why the L1 vowel system would have shifted in just one direction, rather than in two or more directions: raising of the Quichua vowel space was sufficient to accommodate the non-high Spanish vowels at distinct positions in a combined L1/L2 vowel space, such that no lowering (e.g., of Quichua /a/) or fronting/backing was necessary. However, this account does not explain why the direction of the shift was upward, rather than downward. Presumably, L1 vowels shifted upward because there was simply more space between the high and mid vowels than between the mid and low vowels, and possibly also because Spanish /a/ may have existed at an F_1 ceiling, which could have prevented Quichua /a/ from shifting lower than it.

The findings of Flege (1987b) and Guion (2003) are similar in that they both evince L1 phonetic drift of vowels, but they differ in that Flege examined individual L1 and L2 vowels, while Guion examined the L1 and L2 vowel spaces in their entirety. The investigation of the vowel space as a whole allows for a level of analysis that is not possible when only pairs of close vowels are considered. Thus, whereas the forward drift of French /u/ seen in Flege (1987b) was attributed to influence from L2 English /u/ specifically, the upward drift of Quichua /ɪ, a, ʊ/ seen in Guion (2003) was analyzed as a system-wide development motivated by pressures toward vowel dispersion. In this respect, the drift in Quichua vowels and the drift in French /u/ may be said to have opposite motivations. To be specific, the latter example was analyzed as assimilatory, as it brought the realization of French /u/ closer to that of the high-F_2 English /u/; on the other hand, the former example was essentially dissimilatory, as it shifted the Quichua vowels away from the Spanish non-high vowels for the sake of perceptual contrast in a shared L1-L2 phonetic space. The crucial detail of Guion's analysis is that the dissimilation did not occur between any two vowels in particular; rather, it acted as a system-wide force that served to ensure sufficient dispersion between distinct vowel categories with minimal movement. However, in reality, a plain dissimilatory account of the Quichua vowel raising is also possible. One can imagine that Quichua /ɪ, ʊ/ raised to dissimilate from Spanish /e, o/, and that the positions of the low vowels resulted from a pull-and-push chain of events: Spanish /a/ raised in approximation to the higher Quichua /a/, which then raised further to dissimilate from Spanish /a/ (cf. Mack 1990). It is difficult to distinguish between these two analyses due to the small vowel inventory of the L1 in this case. The occurrence of three separate vowel shifts in the same direction is not an unlikely coincidence, whereas the occurrence of ten separate vowel shifts in the same direction would be much harder to believe. Examination of bilingual production in an L1 and L2 with larger vowel inventories would thus make it clearer whether the Quichua vowel raising found by Guion (2003) was a coordinated occurrence or the sum total of three separate shifts that happened to go in the same direction.

Dissimilatory L2 influence on L1 vowels was also documented in work on L1 Italian-L2 English bilinguals conducted by Flege et al. (2003). A series of acoustic analyses in this study examined production of mid front vowels in four groups of bilinguals differing in AOA from Italy to Canada and in frequency of continued L1 use. Crucially, the mid front vowels of the L1 and the L2 here differed in terms of diphthongization: Italian /e/ is monophthongal, whereas English /e/ is diphthongal, realized as [eᶦ]. Results showed that late bilinguals tended to produce English /e/ with less formant movement than native English speakers, regardless of their current level of L1 use. This pattern of undershoot was attributed to failure to establish a new phonetic category for English /e/, which allowed for cross-linguistic influence from monophthongal Italian /e/ via a shared phonetic category. On the other hand, early bilinguals who used Italian infrequently were found to produce English /e/ with significantly more formant movement than even native English speakers. The exaggeration of formant movement in these bilinguals' productions was claimed to follow from the dissimilation of a phonetic category that had been separately established for English /e/ from the phonetic category for Italian /e/.

Age of L2 acquisition and amount of L2 experience were both found to affect the vowel production of L1 Korean-L2 English bilinguals studied by Baker and Trofimovich (2005). Six English vowels (/i/, /ɪ/, /ɛ/, /æ/, /u/, /ʊ/) and five Korean vowels (/i/, /e∼ɛ/, /u/, /ɨ/) were examined, along with four groups of bilinguals crossed by age of acquisition and amount of experience (early-low, early-high, late-low, late-high). Results of acoustic analyses showed that while late bilinguals did not produce Korean vowels differently from monolingual Korean speakers, they produced the English vowels /ɪ/, /ʊ/, and /æ/ differently from monolingual English speakers—namely, in a manner approximating the similar Korean vowels /i/, /u/, and /ɛ/, respectively. Whereas late bilinguals manifested only an effect of L1 on L2, early bilinguals' production showed both L1 influence on L2 as well as L2 influence on L1. Like late bilinguals, early bilinguals produced the English vowels /ɪ/, /ʊ/, and /æ/ differently from English monolinguals, in a manner similarly convergent with nearby Korean vowels. In addition, early-low bilinguals resembled late bilinguals in producing Korean vowels no differently from age-matched monolingual Korean controls. In contrast, early-high bilinguals produced the Korean vowels /i/, /u/, and /ɛ/ differently from Korean monolinguals, in a manner generally suggestive of dissimilation from nearby English vowels. These results suggested that L2 influence on L1 is favored when L2 experience both begins early and accumulates in great quantity.

Thus, an abundance of evidence from studies of bilingual and L2 phonology indicates that L1 production can be profoundly affected by L2 experience. L2 influence on L1 speech has been found in studies of early, late, simultaneous, and sequential bilinguals; studies of children and adults; studies of overall accent; studies of individual segment types such as stops, fricatives, approximants, and vowels; and studies examining a variety of L1s and L2s. A recurrent finding of this literature is that L1 sounds tend to drift toward the closest L2 sounds. However, they may also drift away from L2 sounds in order to maximize contrast within a shared phonological system. Consistent with the predictions of the SLM, dissimilatory phonetic drift of L1 usually occurs in early bilinguals, individuals who received L2 exposure from a young age and were thus more likely to develop separate phonetic categories for L2 sounds. Phonetic drift of both the assimilatory and dissimilatory kind has often been analyzed with respect to individual category pairs, but as seen in the case of Quichua-Spanish bilinguals, it may be possible for L2 to exert an influence on L1 at a more global level.

2.4.4 Summary

Whether L1 is an endangered language or a non-endangered language, L1 production has been overwhelmingly shown to be affected by significant experience in an L2. L2 influence on L1 is pervasive in attrition contexts, as expected. This could be attributed to lack of L1 use or, in the case of heritage speakers and child bilinguals, to incomplete acquisition of the L1. However, L1 phonetic drift occurs even in individuals who reached an adult monolingual-like level of L1 proficiency prior to L2 exposure—that is, late L2 learners. Findings on late L2 learners suggest that while L1 phonetic drift is not inevitable, at least in certain circumstances (e.g., VOT in formal speech: Major 1992; F_2 of /u/ in English: Flege 1987b), reaching a high level of L2 proficiency usually leads to changes in L1 production, and this effect occurs even when L1 continues to be used on a daily basis.

2.5 Motivations and Predictions

As alluded to in Chapter 1, this dissertation has one main goal: to arrive at a better understanding of *how* and *why* L1 phonetic drift occurs. The research reviewed in this chapter has suggested that while a decline in L1 use may contribute to phonetic drift, this is not the main cause. Rather, L2 experience appears to be the primary factor driving changes to L1 production. It remains unclear, however, how much L2 experience is required to result in significant modifications to L1 production, since previous studies have documented this phenomenon almost entirely in individuals with large amounts of L2 experience—either L2-dominant speakers in situations of attrition (e.g., last speakers, heritage speakers) or fluent bilinguals. This bias in the literature accords with what the prevailing view seems to be—namely, that a large amount of L2 experience (as indicated by a high level of L2 proficiency) is required to influence L1 in any significant way:

> ...a L2 that is hardly mastered should not have much influence on L1, while a L2 which is mastered to a high degree should exert more influence. (Major 1992:201)

Although the findings of Yusa et al. (2010) indicate that children with relatively little L2 experience may also show signs of L1 phonetic drift (see also Ward, Sundara, Conboy, and Kuhl 2009 for convergent findings with infants), these results are ambiguous since children also have relatively little L1 experience; hence, L1 phonetic drift here can be attributed to underdeveloped L1 representations that are still maturing. What would improve our understanding of phonetic drift, therefore, is a study that examines individuals with fully developed L1 representations from the onset of L2 exposure. If such a study found phonetic drift even in this case, it would demonstrate a fundamental continuity of language development over the lifespan, both in a temporal and experiential sense: the development of linguistic representations drawn upon in speech production would have to be thought of as constantly being updated by input from a wide range of sources, rather than maturing towards a well-defined endpoint on the basis of data from one and the same language.

This dissertation is that study. It approaches the task of investigating the nature and time course of L1 phonetic drift by starting from the linguistic conditions least likely to produce evidence of drift: adult L1 monolinguals, who are starting to learn an L2 that is both genetically and orthographically unrelated to the L1, producing L1 speech in a formal style (specifically, monolingual speakers of American English, who are starting to learn Korean, producing citation-form English in a word reading task). Will these speakers also show phonetic drift in their L1 production? The principles of the SLM lead us to predict that they will. Recall that cross-language influence in the SLM is based on equivalence classification, a basic cognitive process that aids in perceptual categorization. Although left implicit in the model, the nature of equivalence classification, as a mechanism used in normalization and categorization, is almost certainly automatic. If, consequently, novel L2 sounds are perceptually linked to close L1 sounds automatically, it follows that the potential for cross-language influence will be immediate. Thus, the first hypothesis of this study is that ADULT L2 LEARNERS WILL MANIFEST L1 PHONETIC DRIFT EARLY IN THE COURSE OF L2 ACQUISITION, rather than after a high threshold of L2 experience has been passed. This hypothesis is tested for the two category types for which L1 phonetic drift has been most widely reported: consonant laryngeal categories (Chapter 4) and vowel quality categories (Chapter 5).

Assuming that L1 phonetic drift occurs, the second question to ask is whether the drift will be assimilatory or dissimilatory. The principles of the SLM lead us to predict that phonetic drift in this case will generally be assimilatory. Recall that according to the SLM, dissimilation between L1 and L2 sounds only occurs when the L2 sound was perceived as different enough from the closest L1 sound to have formed its own phonetic category. The SLM also states that new phonetic categories for L2 sounds are established less often with increasing age of L2 learning, which implies that dissimilation is unlikely to happen in adult learners. Consistent with this conclusion, which finds support in the existing literature (Section 2.4.3), the second hypothesis of this study is that ADULT L2 LEARNERS WILL SHOW L1 PHONETIC DRIFT THAT IS ASSIMILATORY to L2 sounds.

The third question to ask is whether the drift will occur at a segmental level or at a non-segmental level (e.g., a natural class of sounds). Only L1-L2 perceptual linkages at the level of the segment (cf. phonetic category) are discussed in the SLM. This is also true of the PAM(-L2), which discusses perceptual assimilation in terms of relations between the gestural constellations of individual segmental categories. As such, it would be logical to expect L1 phonetic drift to occur via segment-based cross-linguistic connections. However, previous findings—in particular, those of Nielsen (2008) and Guion (2003)—lead us to different predictions. The results of Nielsen (2008) for VOT imitation, which showed talkers being influenced by exposure to model speech at a level generalizing to all stops with the same laryngeal specification, suggest that, at least in the domain of VOT, L1 representations change at the level of a relevant natural class. What this means is that hearing heavily aspirated tokens of /p^{h}/, for example, is expected to affect VOT production not just for the segment /p^{h}/, but for the natural class of all stops with the same specification for the relevant features (i.e., all voiceless aspirated stops), because the retuning of production happens at the level of an 'aspiration' feature (e.g., [+spread glottis]), rather than for individual segments. Thus, the third hypothesis in this study is that, rather than occurring between specific consonants, L1 PHONETIC DRIFT IN VOT WILL OCCUR AT THE LEVEL OF THE NATURAL CLASS, affecting all stops of the same laryngeal type to the same degree. In a similar way, the results of Guion (2003) for vocalic drift, which showed three L1 vowels moving in concert, suggest that L1 representations for vowels change not at a segmental level, but at a systemic level. It is assumed that the similar patterns of drift found for these three L1 vowels did not result from coincidence, but from a system-level shift. Thus, the fourth hypothesis in this study is that, rather than being realized differently for each vowel, L1 PHONETIC DRIFT IN VOWEL FORMANTS WILL OCCUR AT THE LEVEL OF THE VOWEL SYSTEM, resulting in a global shift affecting all L1 vowels in a similar manner.

In Chapter 3, the design and methods of the current longitudinal study are presented, including details about the participants and analysis protocols for the acoustic measurements discussed in Chapters 4–5. An introduction to the L2 sounds serving as triggers for L1 phonetic drift is provided in each of the chapters discussing the experiments—background on Korean stop consonants in Chapter 4, and background on Korean vowels in Chapter 5.

Chapter 3

Methodological Overview

3.1 Introduction

This chapter describes the overall design of the study, the characteristics of the population under examination, and the methods used in the analysis of the data collected. The study population consists of late learners of Korean undergoing an intensive program of Korean language training in South Korea. Over the course of this program, learners participated in two production experiments. Experiment 1K investigated their production of Korean, the target language (L2), while Experiment 2E investigated their production of English, the native language (L1).[1]

3.2 Participants

The population of interest comprises adult L1 English-speaking L2 learners of Korean. A sample of 40 learners participated in the study, with four learners dropping out in the middle and seventeen others being excluded for other reasons: six because of childhood exposure to Korean, four because of significant prior study, four because of regular exposure to and/or use of another language at home besides English, two because of significant hearing and/or speech impairments in childhood, and one because of failure to follow directions in Week 1 of the study. These learners studied other languages (mostly Spanish and French) in the course of their formal education; however, in the background questionnaire they completed prior to their participation in the study, they reported English to be their native language, their best language, and the language used at home and no regular communicative use of another language besides English. Moreover, they reported no significant prior exposure to Korean, prior study of Korean, or prior travel to Korea lasting longer

[1]Learners also participated in experiments examining their non-linguistic perception, their perception of the L2, and their ability to imitate the L2. Some preliminary results are reported in Chang (2009c).

Participant	Gender	Age (yrs.)	Home state	Other languages (duration of study)
LF03	female	21	TX	Spanish (8 yrs.), French (1 yr.)
LF05	female	21	MA	Spanish (12 yrs.)
LF06	female	23	WA	French (4 yrs.), Spanish (1 yr.)
LM13	male	22	OR	Spanish (4 yrs.), Latin (2 yrs.), Ancient Greek (2 yrs.), Modern Greek (4 mos.)
LF16	female	22	FL	Spanish (3 yrs.)
LF18	female	26	MO	French (6 yrs.), Latin (6 yrs.), Italian (1 yr.)
LF19	female	22	OH	French (6 yrs.), German (1 yr.)
LM23	male	22	IL	Hebrew (13 yrs.), Setswana (4.5 mos.)
LF25	female	22	MA	French (6 yrs.), Spanish (1 yr.)
LF28	female	22	IN	French (9 yrs.)
LF29	female	22	WA	Hebrew (15 yrs.), Spanish (3 yrs.)
LF31	female	22	NJ	Spanish (6 yrs.)
LF32	female	22	TX	Spanish (15 yrs.)
LF37	female	22	PA	French (9 yrs.)
LM44	male	22	IN	Spanish (3 yrs.)
LF46	female	21	IL	German (8 yrs.)
LF47	female	22	RI	Spanish (6 yrs.), Latin (2 yrs.)
LF54	female	22	IA	Spanish (7 yrs.), French (2 yrs.)
LF55	female	22	NY	French (6 yrs.)

Table 3.1: Demographic data, L2 learners of Korean.

than one week. Thus, in the end the group of L2 learners discussed here contains 19 "functionally monolingual" (in the sense of Fishman 1972:141, Baker and Jones 1998:158, and Best and Tyler 2007:16) native speakers of American English (sixteen females, three males). All were paid for their participation. Their mean age was 22.1 years (range of 21–26), and they came from hometowns dispersed across a variety of U.S. dialect regions, as summarized in Table 3.1.

In addition to the group of L2 learners, a control group of nine native Korean speakers (seven females, two males) participated in the production experiments. All were paid for their participation. Their mean age was 27.8 years (range of 22–34). These nine Korean speakers provided most of the learners' L2 input. Participants NF1, NF2, NF3, NM4, NF5, NF6, and NF7 were teachers in the Korean language

Participant	Gender	Age (yrs.)	Home province
NF1	female	28	Jeju
NF2	female	23	N. Jeolla
NF3	female	27	Seoul
NM4	male	31	Gyeonggi
NF5	female	31	Seoul
NF6	female	34	Seoul
NF7	female	31	Seoul
NM8	male	23	N. Chungcheong
NF9	female	22	Gangwon

Table 3.2: Demographic data, native speakers of Korean.

program in which learners were enrolled (described in Section 3.3), while participants NM8 and NF9 were resident assistants in the dormitory where learners were living during the language program. Having been educated in South Korea, where formal English instruction is compulsory from as early as primary school, these native Korean speakers had all received some degree of schooling in English, but they were strongly dominant in Korean at the time of this study and can be considered representative of young Korean speakers in contemporary South Korea. All the teachers spoke Standard Korean, having been trained in Seoul; five of the seven also hailed originally from Seoul or the surrounding Gyeonggi province, as shown in Table 3.2.

3.3 L2 Exposure

At the time of data collection, learners were living on the campus of a South Korean university and embarking on a six-week course of intensive Korean language instruction in preparation for a year of cultural immersion and exchange. They were divided into five beginner-level classes, each taught by two instructors from the group of native Korean speakers shown in Table 3.2. On average learners had four hours of class a day, for a total of approximately 82 hours of instruction by the end of the program (roughly equivalent in content to one semester of college-level Korean). The structure of the language program is presented in Table 3.3, while the class affiliations of the learners and their instructors are summarized in Table 3.4. Beginning-level instructors, who were all female, varied in terms of teaching style and the amount of English they used in class; however, all five classes were conducted in Korean the vast majority of the time, and instructors followed the same general curriculum using the same main textbooks.

In exit questionnaires, most learners reported that time spent in class (typically between 12 and 20 hours/week) constituted the majority of the time they heard

Week	Content
1	placement test & interview (2 hrs.), classes (12 hrs.)
2	classes (12 hrs.)
3	classes (15 hrs.)
4	classes (20 hrs.)
5	classes (20 hrs.)
6	final exams (4 hrs.), classes (3 hrs.), final presentations (2 hrs.)

Table 3.3: Structure of the Korean language program.

Class	Instructor 1	Instructor 2	Learners
Beginner 1	NF5	NF1	LF06, LF25, LF46
Beginner 2	NF7	NF5	LF03, LF05, LM13, LF29
Beginner 3	NF1	NF7	LM23, LF54, LF55
Beginner 4	NF2	NF6	LF16, LF18, LF19, LF31, LM44, LF47
Beginner 5	NF6	NF2	LF28, LF32, LF37

Table 3.4: Class affiliations of learners and instructors.

Korean during the six weeks of the study. Learners received some additional exposure to Korean in instructors' office hours, during any extracurricular activities they were involved in (e.g., *taekwondo*, *hanji*, drumming, calligraphy, the university English club), in engaging the resident assistants in their dormitory, and in the course of exploring the university campus and surrounding town. Most learners estimated the amount of this additional exposure to vary between three and ten hours/week, though three learners estimated it to be closer to 20 hours/week. What is consistent about their descriptions is that this additional exposure to Korean was short and/or sporadic in comparison to the long, consistent exposure received during class time (which learners were likely attending to in a different way). For instance, although *taekwondo* classes took up four hours/week and were taught by Korean speakers, comparatively little language was used during this time; moreover, the language that was used comprised mostly numbers and a fixed set of terms for physical maneuvers. Learners' judgments of the amount of time they spent speaking (rather than simply hearing) Korean outside of class show less variation, with estimates varying between one and six hours/week. Thus, the vast majority of the Korean they produced during the time period of the program was within the context of their classes.

The type of language learning situation in which the learners found themselves was, therefore, a cross between typical Second Language Acquisition (SLA), in which learners are immersed in an L2 environment and acquire the L2 largely "in the wild", and typical Foreign Language Acquisition (FLA), in which learners study the L2 formally in an L1 environment. In the current study, learners were living in Korea, but receiving most of their L2 exposure via structured formal instruction in a classroom setting where all their fellow learners shared the same L1 background. After class, with the exception of a few extracurricular activities learners operated predominantly in L1. For this reason, the L2 exposure they were receiving cannot be considered traditional SLA immersion. Rather, it is best thought of as intermediate in intensity between SLA and FLA.

3.4 Experiments

Experiments 1K and 2E were production experiments examining learners' pronunciation of L1 and L2 sound categories. In both experiments, production of the sounds of interest was elicited via a reading task in which participants were given orthographic cues to the items they were to produce. Experiment 1K elicited production of L2 (Korean), while Experiment 2E elicited production of L1 (English).

Both experiments were longitudinal in nature and run a total of five times, each time in the space of 48 hours between the end of one week of instruction and the beginning of the following week of instruction. In this way, the amount of Korean instruction received prior to participation in each experiment was kept equal across participants.

Experiments 1K and 2E were almost always completed in one session, with a break between experiments. Since Experiment 1K was preceded by another experiment involving a task in Korean, Experiment 1K was completed before Experiment 2E in order to require only one switch between languages (see, e.g., Grosjean 2001). This order was furthermore kept constant across participants so as to allow for both cross-sectional and longitudinal comparisons.

3.5 Procedure

The experiments were generally run in a quiet room in the dormitory where learners were living during the language program (the only exception being after Week 2, when the experiments were run in the hotel where learners were staying during a weekend trip). In these experiments, stimuli were presented and responses recorded in DMDX (Forster 2008) on a Sony Vaio PCG-TR5L laptop computer. Participants recorded their responses using an AKG C420/520 head-mounted condenser microphone, which was connected either to the computer via an M-AUDIO USB preamp or to a Marantz PMD660 solid-state recording device. In both cases, audio was recorded at 44.1 kHz and 16 bps. A representative picture of the experimental environment is shown in Figure 3.1.

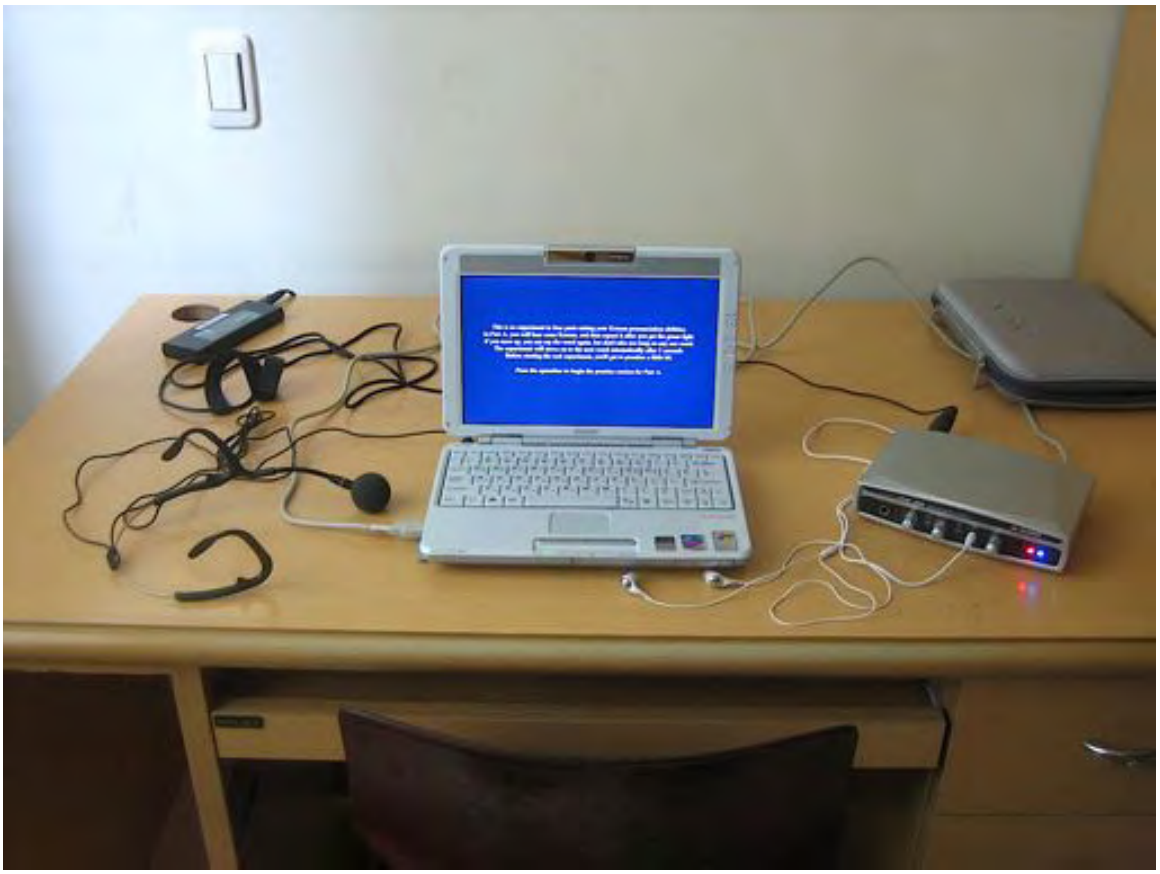


Figure 3.1: Experimental environment for Experiments 1K and 2E.

In addition to the production experiments, learners completed two questionnaires as part of their participation in the study. The first was a detailed entrance questionnaire about learners' linguistic and social background (Appendix A), which they submitted prior to their first completion of Experiment 1K. In the entrance questionnaire, learners were asked about geographic and social affiliations, previous experience with Korean and Korea, their home environment, education, language competence, motivation, and learning goals. The second questionnaire was a shorter exit questionnaire about learners' experiences in the language program and in the study (Appendix B), which they submitted during Week 6 of the language program. In this questionnaire, learners were asked about the particular dynamics of their Korean class; the amount of time they studied, used, and were otherwise exposed to Korean during the language program; their final level of Korean proficiency, level of motivation, and learning goals; their ideas about what the objectives of the study were; and strategies they might have used in the experiments.

3.6 Stimuli

The set of stimuli consisted of 22 Korean and 23 English monosyllables representing most of the phonemic contrasts in the two languages. Members of a subgroup of stimuli were maximally similar to each other in segmental makeup (e.g., Korean 후 /hu/ vs. English *who'd* /hud/) for the purposes of cross-language acoustic comparisons. English monosyllables were generally of the form CVC to allow for lax vowels, while Korean monosyllables were generally of the form CV to make them easier for novice learners to read. The same set of stimuli was used in every week of the study (see Table 3.5 for a full list[2]).

Each subset of stimuli listed in Table 3.5 was meant to test one or more of the hypotheses formulated in Chapter 2. The items beginning with stop consonants were meant to test for phonetic drift in voice onset time (specifically, the hypothesis that drift in voice onset time would be subphonemic and occur at the level of the natural class). Thus, stop-initial items were included to elicit productions of stop consonants that could be measured for voice onset time (Section 3.7.1), the primary distinguishing characteristic of different stop voicing (laryngeal) categories, as well as for fundamental frequency onset in the following vowel (Section 3.7.2). The items beginning with /h/ were meant to test for phonetic drift in vowels (specifically, the hypothesis that phonetic drift in vowel production would be systemic and occur over the whole vowel space). Thus, items beginning with onsets having no oral place of articulation were included so as to elicit productions of vowels that could be measured for formant frequencies (Section 3.7.3), the primary distinguishing characteristic of

[2] IPA transcriptions use the extended IPA symbols for weaker and stronger articulations (International Phonetic Association 1999:189) to transcribe lenis obstruents (e.g., /t͉/) and fortis obstruents (e.g., /t͈/).

Table 3.5: Korean and English stimuli used in Experiments 1K and 2E.

Korean		English		Korean		English	
바	/p͉a/	*'bot*	/bɑt/	히	/hi/	*heed*	/hid/
빠	/p͈a/					*hid*	/hɪd/
파	/pʰa/	*pot*	/pɑt/	해	/hɛ/	*hate*	/het/
다	/t͉a/	*dot*	/dɑt/			*head*	/hɛd/
따	/t͈a/					*had*	/hæd/
타	/tʰa/	*tot*	/tɑt/	호	/ho/	*hoed*	/hod/
가	/k͉a/	*got*	/gɑt/	후	/hu/	*who'd*	/hud/
까	/k͈a/			흐	/hɨ/	*hood*	/hʊd/
카	/kʰa/	*cot*	/kɑt/	허	/hʌ/	*hut*	/hʌt/
						hawk	/hɔk/
사	/s͉a/	*sod*	/sɑd/	외	/wɛ/	*wait*	/wet/
싸	/s͈a/	*shot*	/ʃɑt/			*wet*	/wɛt/
시	/s͉i/	*seed*	/sid/	위	/wi/	*wee*	/wi/
씨	/s͈i/	*sheet*	/ʃit/	알	/aɭ/	*all*	/ɔl/

different vowel qualities, with minimal coarticulatory influence from onset consonants. The remaining stimuli comprised control and filler items.

3.7 Acoustic Analysis

The acoustic data from recordings comprised measurements of voice onset time (VOT) in word-initial plosives, fundamental frequency (f_0) at the onset of the following vowels, and first formant frequency (F_1) and second formant frequency (F_2) in vowels. All acoustic measurements were taken manually in Praat (Boersma and Weenink 2008) on a wide-band Fourier spectrogram with a Gaussian window shape (window length of 5 ms, dynamic range of 50 dB, pre-emphasis of 6.0 dB/oct) or the corresponding waveform. Data visualization and statistical analyses were conducted in R (R Development Core Team 2010).

3.7.1 Voice Onset Time

VOT was defined as the time at voicing onset minus the time at plosive release (the beginning of the burst interval). Thus, VOT was negative when the voicing onset preceded the plosive burst ("prevoiced" stops) and positive when the voicing onset followed the plosive burst ("lag-voiced" stops). It should be noted that by a "prevoiced" stop, what is meant is not a stop with voicing preceding the stop closure, but a stop with voicing during the stop closure (before the burst). The point of

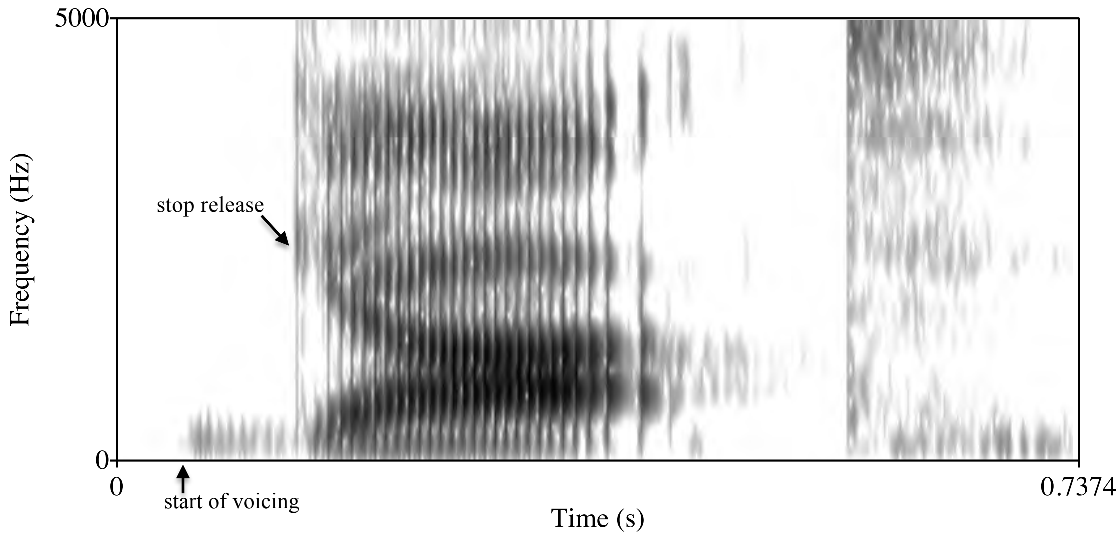


Figure 3.2: Marking of negative VOT in an initial prevoiced plosive (*got*, Token 2, LM13, Week 3).

voicing onset was taken to be the first point at which a voicing bar with clear glottal striations appeared in the spectrogram. In unclear cases, this point was marked early (i.e., at the leftmost point of an ambiguous interval in the spectrogram such as breathy phonation, which often showed a voicing bar similar to modal phonation).

Examples of how prevoiced and lag-voiced plosives were marked for VOT are shown in Figures 3.2 and 3.3. Note that in prevoiced plosives, voicing often died out well before the burst, particular when it began more than 250 ms beforehand. However, the negative VOT values do not distinguish between prevoiced plosives which were fully voiced until the burst and those which were not; they represent only the time at which voicing began prior to the burst.

3.7.2 Fundamental Frequency

In order to obtain stable f_0 measurements, f_0 onset was measured over three glottal periods. An interval of three periods was chosen, rather than just one period (or one time point in an automated f_0 analysis such as autocorrelation), due to the often large disparities between the wavelength of the first period and the wavelengths of the following periods, which would have introduced spuriously high and low f_0 estimates into the data. Thus, f_0 onset was measured by taking the combined wavelength of the first three regular glottal periods in the vowel and converting this into a frequency value (by inverting and then multiplying by 3). The three-period interval was marked off on the waveform, with an initial period generally being skipped if it

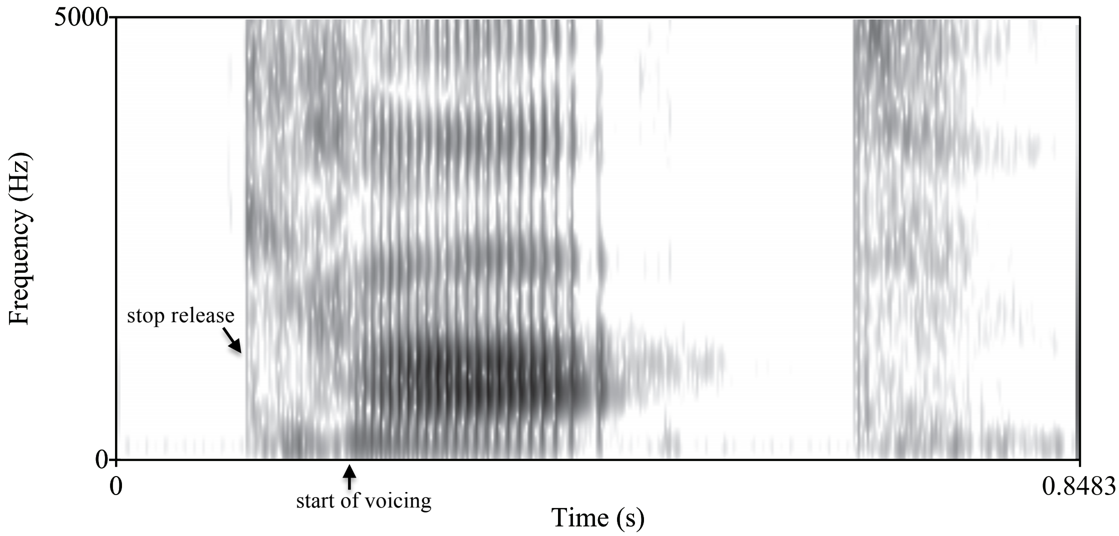


Figure 3.3: Marking of positive VOT in an initial lag-voiced plosive (*cot*, Token 2, LM13, Week 3).

was more than 33% longer or shorter than the following period. In unclear cases, the beginning of the interval was marked early (i.e., a slightly irregular initial period was included in the interval) so as to obtain a measurement as close to the onset of the vowel as possible. No f_0 measurement included in the final data set resulted from a three-period interval that skipped over more than five periods; if a token being analyzed was such that the earliest interval of three regular periods occurred more than five periods into the vowel (i.e., the vowel started off with an extended interval of creaky or irregular phonation), the token was discarded.[3] Figure 3.4 shows an example of a three-period interval marked right at the beginning of a vowel with regular phonation from onset, while Figure 3.5 shows a three-period interval marked a few periods into the vowel so as to skip over the initial irregular periods.

In order to put male and female learners on the same f_0 scale, raw f_0 measurements were standardized to z-scores by learner, by subtracting the learner's mean f_0 over the entire duration of the study and dividing by the square root of the learner's variance in f_0 over the entire duration of the study.

3.7.3 Vowel Formants

Measurements of vowel formants were extracted from spectrograms annotated for vowel onset and offset by the author and two additional researchers, Daiana Chang

[3]Only a small percentage of tokens (between 1.9% and 5.0% depending on the language) were discarded by this criterion or for other reasons.

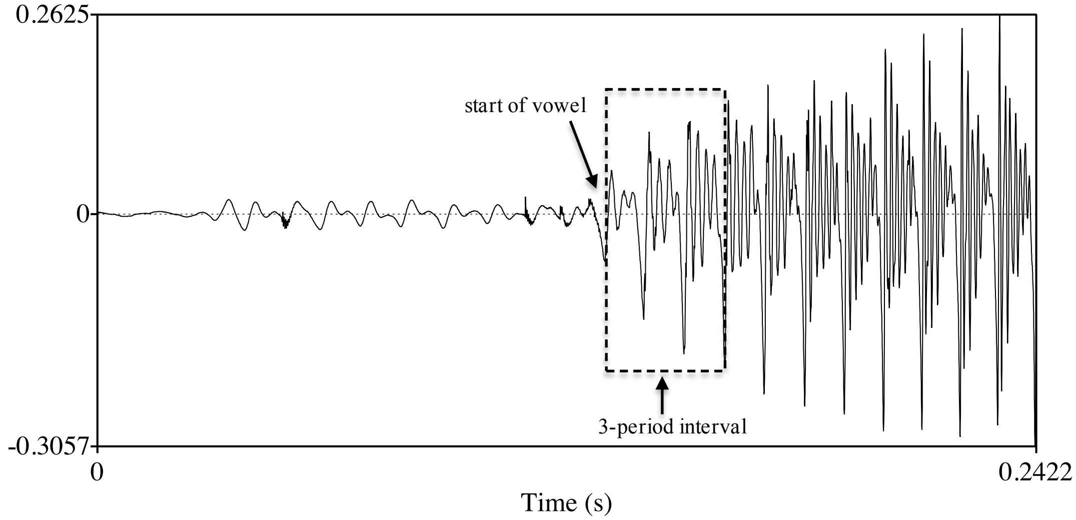


Figure 3.4: Marking of a three-period interval at the beginning of a vowel with regular phonation (*got*, Token 1, LM13, Week 3).

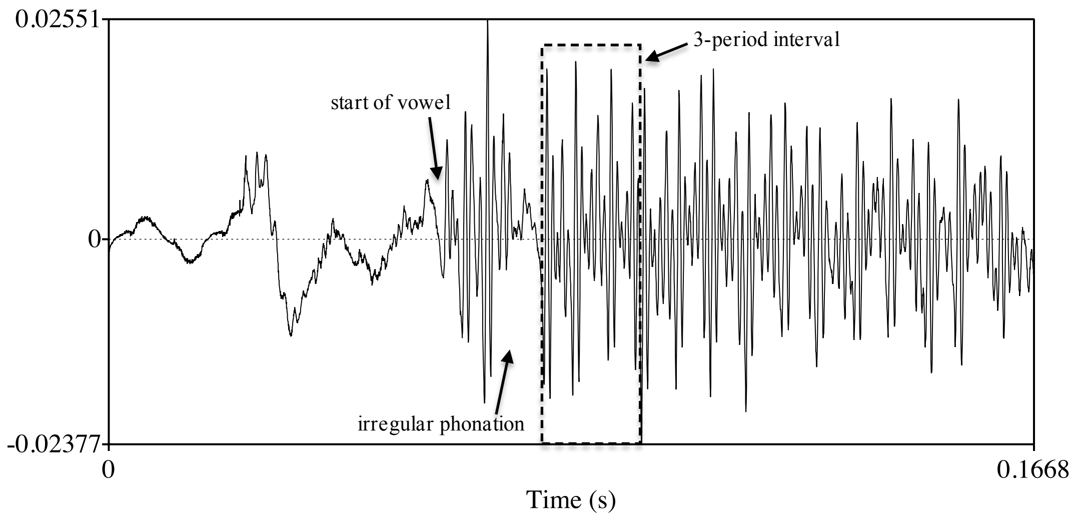


Figure 3.5: Marking of a three-period interval at the beginning of a vowel with some irregular phonation (*'bot*, Token 3, LF05, Week 3).

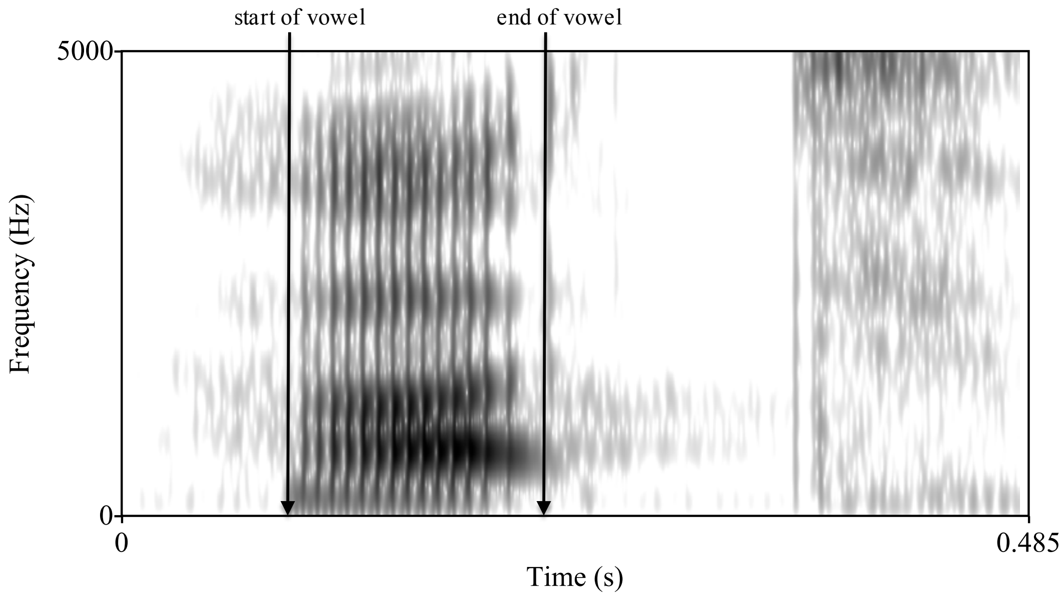


Figure 3.6: Marking of vowel onset and offset in a consonant-final word (*hut*, Token 1, LM13, Week 3).

and Kevin Sitek, who were trained to use the same annotation protocols. The beginning of the vowel was marked at the first glottal striation. In unclear cases, this point was marked early (i.e., at the leftmost point of an ambiguous interval) so as not to inadvertently exclude speech. The end of the vowel was marked, in words with a following coda consonant, at the final glottal striation (Figure 3.6), and in words with no following consonant, at the point where a clear F_1 and F_2 were no longer visible (Figure 3.7). In unclear cases, this point was marked late (i.e., at the rightmost point of an ambiguous interval), again to avoid inadvertently excluding speech.

Mean values of F_1 and F_2 were measured over the middle 50 ms of the vowel interval demarcated in this way. The analysis method was linear predictive coding, using the Burg algorithm (Childers 1978) in Praat. The frequency range and number of formants entered into the formant analysis were obtained by looking at a few spectrograms from the given participant and adjusting the default parameters until tracking of F_1 and F_2 was smooth and accurately followed the formant bands visible in the spectrogram. The frequency range usually went up to a value between 5000 Hz and 6000 Hz, and the number of formants typically fell between 4.5 and 6.0.

Following this automatic formant extraction, the data were inspected for outliers by vowel and formant, and potential errors were flagged. Spectrograms of all tokens were then individually inspected to check that the formant tracking was accurate. When the formant tracking was irregular or inaccurate, the analysis parameters were adjusted until tracking was smooth, and new measurements were extracted. If the

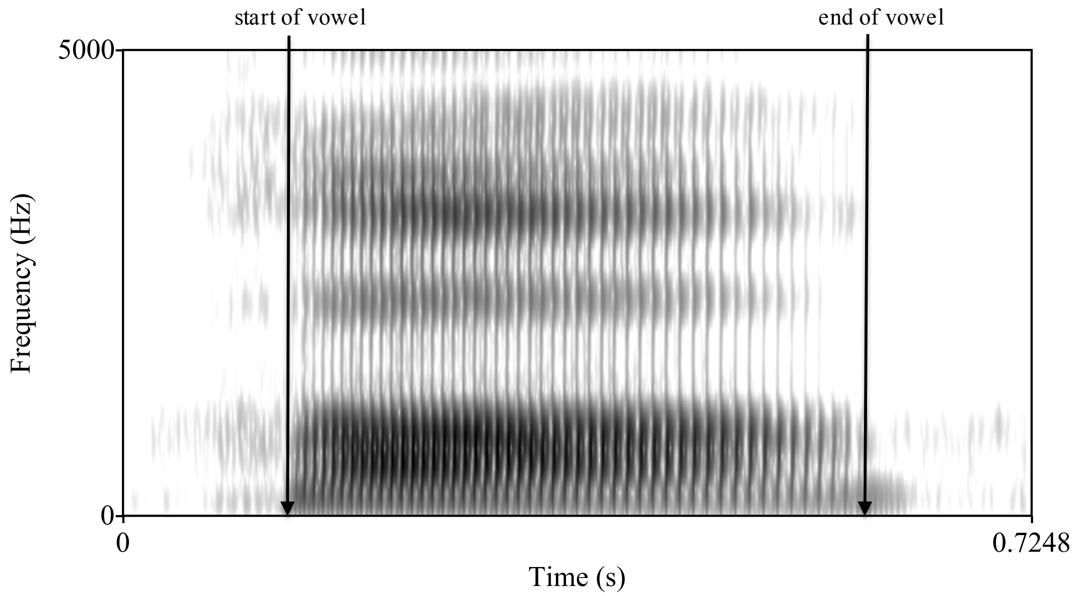


Figure 3.7: Marking of vowel onset and offset in a vowel-final word (허 /hʌ/, Token 1, LM13, Week 3).

formant tracking could not be made satisfactory via adjustment of the analysis parameters, then measurements were taken manually on an average spectrum of the middle 50 ms of the vowel.

3.7.4 Measurement Reliability

Tests of measurement consistency were conducted to check that the data collected were reliable. First, all the VOT and f_0 measurements taken by the author on recordings gathered during Week 3 of testing (approximately 1,600 measurements each of VOT and f_0) were repeated six months after the original measurements were taken, and these two rounds of measurements were compared against each other. This comparison showed very close correspondence between the two rounds of measurements: the average difference between paired VOT measurements was 3 ms, while that between paired f_0 measurements was 4 Hz. In addition, a random set of 180 formant measurements was redone and compared to the corresponding initial formant measurements. This comparison also showed close correspondence between the two rounds of measurements: the average difference between paired F_1 measurements was 10 Hz, and that between paired F_2 measurements was 25 Hz. Thus, the acoustic data appear to be highly reliable.

Chapter 4

Phonetic Drift in Consonants

4.1 Introduction

This chapter reports the results of acoustic analysis on study participants' productions of Korean and English consonants. The focus here is on production of stop manner categories (i.e., "voicing" or laryngeal categories) in utterance-initial position, as these constitute the type of consonantal contrast that has been most widely discussed in the L2 speech literature. The contrasts examined are a three-way contrast in Korean among *lenis*, *fortis*, and *aspirated* plosives and a two-way contrast in English between *voiced* and *voiceless* plosives.

4.1.1 Cross-Linguistic Differences

The three-way Korean laryngeal contrast has been the subject of a great deal of linguistic research. Previous studies have demonstrated that the three laryngeal series differ from each other in domain-initial position along a number of articulatory, aerodynamic, and acoustic dimensions, including linguopalatal contact (Cho and Keating 2001), glottal configuration (Kim 1970; Kagaya 1974), subglottal and intraoral pressure (Dart 1987), laryngeal and supralaryngeal articulatory tension (Kim 1965; Hardcastle 1973; Hirose, Lee, and Ushijima 1974; Dart 1987), intensity at vowel onset (Han and Weitzman 1970), and voice quality at vowel onset (Abberton 1972; Han 1998; Kim and Duanmu 2004).

The two acoustic dimensions most often noted as cues to the Korean laryngeal contrast are voice onset time (VOT) and fundamental frequency (f_0) at vowel onset. Because they are identified with the highest degree of consensus in the literature as significant cues to the Korean laryngeal contrast in both perception and production, these two cues are the focus of this case study of phonetic drift. VOT has been shown to increase going from fortis to lenis to aspirated stops (Lisker and Abramson 1964; Kim 1965; Han and Weitzman 1970; Hardcastle 1973; Kagaya 1974; Hirose et al. 1974; Jun 1993; Kim 1994; Han 1996; Ahn 1999; Lee and Jung 2000; Kim, Beddor,

Table 4.1: Native VOT norms (in ms) for plosives in Korean and English. Korean figures are averaged over the nine native speaker participants, English figures are from Lisker and Abramson (1964:394).

	short-lag		long-lag		
Place of articulation	Korean fortis	English voiced	Korean lenis	Korean aspirated	English voiceless
labial	8	1	56	84	58
coronal	10	5	56	82	70
dorsal	20	21	73	110	80
AVERAGE	13	9	62	92	69

and Horrocks 2002; Cho, Jun, and Ladefoged 2002; Kim 2004; Silva 2006a; Kang and Guion 2008), while f_0 onset has been shown to increase going from lenis to fortis to aspirated stops (Han and Weitzman 1970; Hardcastle 1973; Kagaya 1974; Han 1996; Ahn 1999; Lee and Jung 2000; Kim et al. 2002; Cho et al. 2002; Kim 2004; Silva 2006a; Kang and Guion 2008). The phonetic implementation of the Korean contrast, however, has been undergoing a change in the language, with younger speakers increasingly relying on f_0 to distinguish categories that used to be more distinguishable on the basis of VOT alone (Silva 2006a,b; Kang and Guion 2008). In each of these dimensions there is now considerable overlap between categories, such that VOT and f_0 are both necessary cues for making a full three-way contrast, schematized in Figure 4.1 on the basis of the perception and production data of Kim (2004). The stop production of the native Korean participants in the current study is consistent with this general phonetic space, as shown in Figure 4.2. Thus, in the Korean speech to which learner participants had the most exposure, the Korean laryngeal categories can be assumed to be realized as follows: fortis stops, with short-lag VOT and relatively high f_0 onset; lenis stops, with medium- to long-lag VOT and relatively low f_0 onset; and aspirated stops, with the longest VOT and the highest f_0 onset.

In contrast to the necessary use of VOT and f_0 in making a full three-way contrast among Korean laryngeal categories, VOT alone largely suffices to make the two-way contrast between English laryngeal categories: voiceless stops are characterized by consistently longer VOTs than voiced stops (Lisker and Abramson 1964). The English stops, however, differ in terms of similarity in VOT to Korean stops, as shown in Table 4.1. With regard to the short-lag categories, the VOTs of English voiced stops are very close to those of Korean fortis stops. Compared to the VOTs of English voiced stops, the VOTs of Korean fortis stops are slightly longer in the labial and coronal regions, but virtually identical in the dorsal region. Both overall and at each place of articulation, the difference in VOT between Korean fortis stops and English

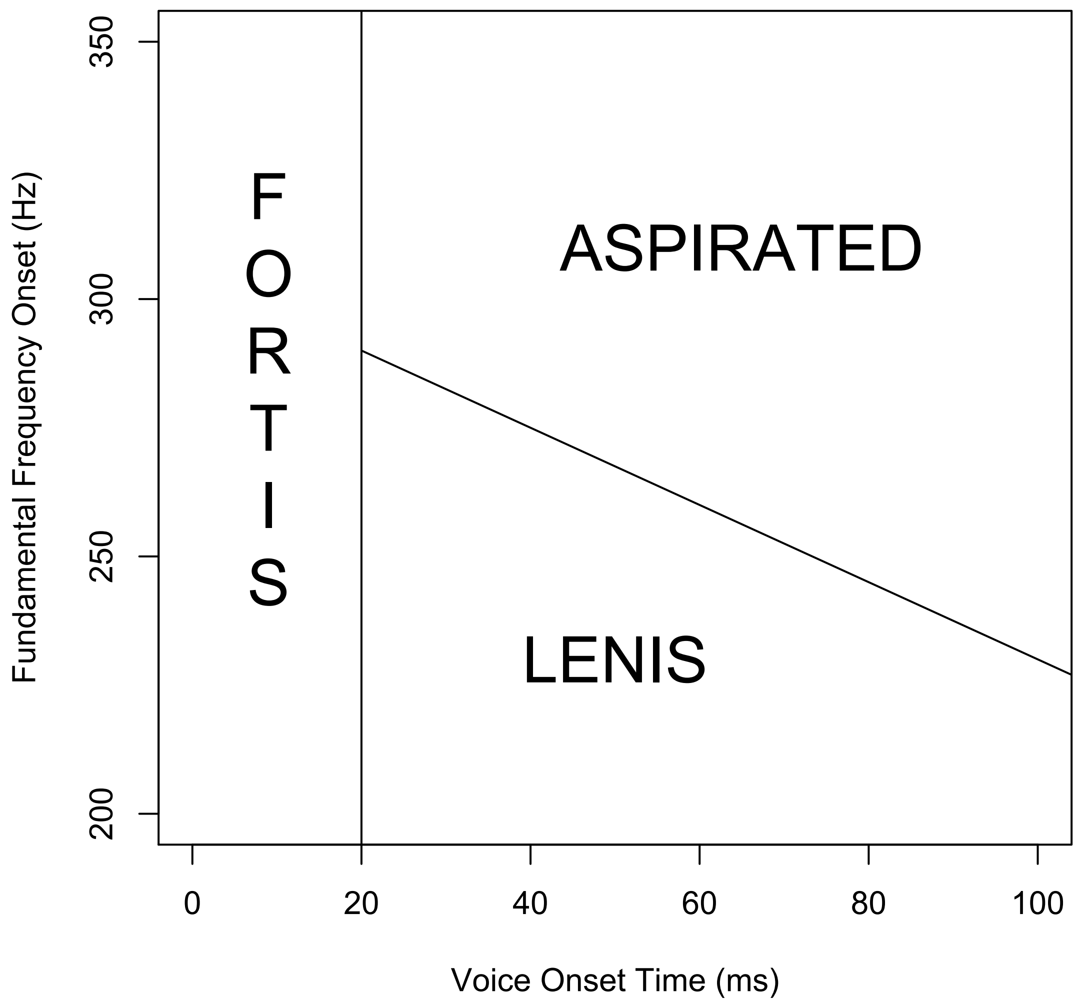


Figure 4.1: Schematic of f_0 onset by VOT in the three-way Korean laryngeal contrast among lenis, fortis, and aspirated plosives and affricates (based on Kim 2004).

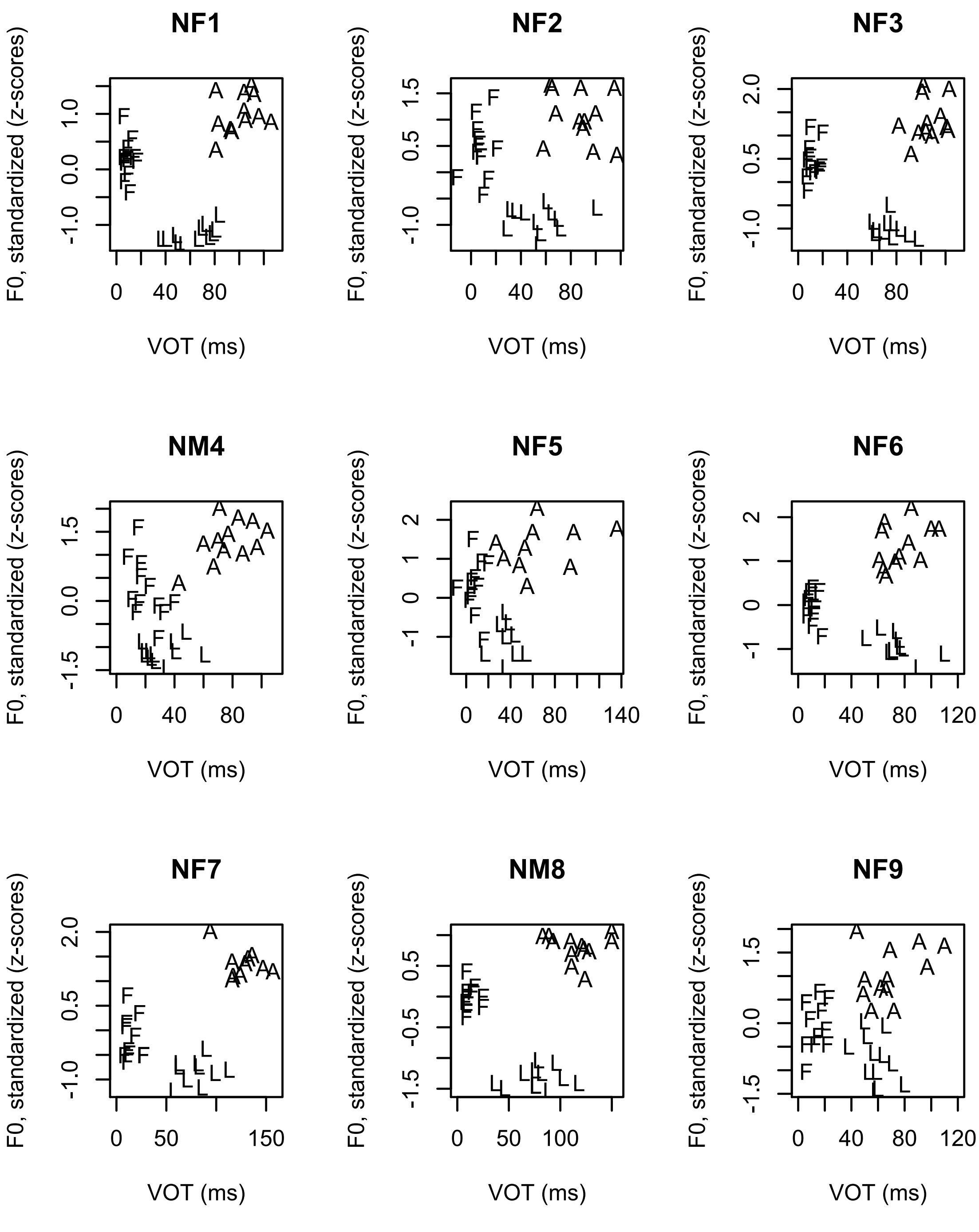


Figure 4.2: Scatterplots of f_0 onset by VOT in native-speaker productions of Korean lenis (L), fortis (F), and aspirated (A) plosives.

voiced stops falls well below the just-noticeable difference (JND) for VOT (Hazan, Messaoud-Golusi, Rosen, Nouwens, and Shakespeare 2009). With regard to the long-lag categories, the VOTs of English voiceless stops are close to those of Korean lenis stops, but less close to those of Korean aspirated stops. Compared to the VOTs of English voiceless stops, the VOTs of Korean lenis stops are shorter at every place of articulation, although the differences again fall under the JND for VOT. On the other hand, the VOTs of Korean aspirated stops are longer at every place of articulation, and in the case of the labials and dorsals (but not the coronals), the differences far exceed the JND for VOT. Thus, it would be reasonable to predict that—on the basis of VOT alone—Korean fortis stops would not be distinguished from English voiced stops, nor Korean lenis stops from English voiceless stops. Korean aspirated stops, however, would likely be distinguished from English voiceless stops due to VOTs that are substantially longer overall.

Though English voiced and voiceless stops are distinguishable in terms of VOT, these categories also differ with respect to the f_0 onset in the following vowel: f_0 starts off lower on average following voiced stops than following voiceless stops (Haggard, Ambler, and Callow 1970; Hombert 1978). Estimates of the magnitude of the f_0 difference between English voiced and voiceless stops vary from study to study: from 4–7 Hz (House and Fairbanks 1953), to 10–13 Hz (Lehiste and Peterson 1961), to 15–20 Hz (Hombert 1978). However, they consistently show the f_0 difference between English laryngeal categories to be more subtle than the f_0 differences between Korean laryngeal categories. In comparison to an f_0 difference between English voiced and voiceless stops that may approach 15–20 Hz, the f_0 difference between Korean lenis and aspirated stops, for example, is much larger, averaging 57 Hz (standard deviation of 14, range of 34–77) for the speakers shown in Figure 4.2. In fact, this sort of pronounced f_0 difference between the Korean lenis and aspirated stops (combined with convergence in VOT ranges) has prompted some researchers to argue that modern Standard Korean has developed tonal contrast (e.g., Silva 2006a). Thus, although direct comparisons of phonetic norms for f_0 are not possible on the basis of what has been reported in the literature, comparisons of f_0 differences suggest that the f_0 onset typical of both English laryngeal categories is substantially lower than the elevated f_0 onset typical of the two laryngeally marked Korean categories (fortis and aspirated). Furthermore, the cross-linguistic differences in f_0 are likely to be noticeable in light of JNDs that have been reported for frequencies in the range of f_0 (Roederer 1973; Harrison 1996).

In perceptual judgments of cross-linguistic similarity between the Korean and English categories, VOT and f_0 both seem to play a role. Given that VOT has been argued to play the primary role in the acquisition of L2 Korean laryngeal categories by L1 English late learners (Kim and Lotto 2002; Shin 2007), the cross-language linkages that are established by learner participants in the current study might be predicted to be between L1 English voiced stops and L2 Korean fortis stops and between L1 English voiceless stops and L2 Korean lenis stops, since these are the pairs of categories that are most similar to each other in VOT (Table 4.1). However, cross-linguistic perceptual data collected by Schmidt (2007), while consistent with the former linkage, are somewhat inconsistent with the latter one. In Schmidt's study, subjects—also L1 English speakers with no prior knowledge of Korean—labeled Korean sounds as the perceptually closest English sound and rated the similarity of the English sound to the Korean sound. Results showed that subjects overwhelmingly labeled Korean lenis stops and aspirated stops as English voiceless stops and Korean fortis stops as English voiced stops, but with different degrees of cross-linguistic similarity: Korean aspirated stops were rated as more similar to English stops than Korean lenis or fortis stops were. Thus, despite Korean lenis stops' closer proximity to English voiceless stops in VOT, Korean aspirated stops were rated as more similar to English voiceless stops, presumably due to greater similarity in f_0. These results suggest that for English-speaking learners of Korean, the "default" equivalence classifications of L2 Korean stops with L1 English stops are of Korean aspirated stops and lenis stops with English voiceless stops and of Korean fortis stops with English voiced stops. However, Korean aspirated stops, as the perceptually more similar category to English voiceless stops, are more likely to be perceptually linked to English voiceless stops than Korean lenis stops are.

4.1.2 Predictions

Chapter 2 presented three hypotheses relevant to this case study. First, it was hypothesized that L1 phonetic drift would occur early in L2 acquisition because the cross-language linkages on which phonetic drift is based are formed at the onset of L2 experience, allowing accruing L2 phonetic input to affect L1 representations from the very first stages of L2 learning. Second, it was hypothesized that, in this study, L1 phonetic drift would be assimilatory to L2, since the participants were late- rather than early-onset L2 learners. Finally, it was hypothesized that, with respect to VOT specifically, L1 phonetic drift would occur at a subphonemic level generalizing across segments (i.e., at the level of the laryngeal natural class), similar to the way in which spontaneous imitation of VOT seems to occur in L1 speech accommodation (Nielsen 2007a,b, 2008).

Given the similarities and differences between the Korean and English laryngeal categories described above, these hypotheses lead to three predictions regarding phonetic drift in the VOT of English voiced and voiceless stops. Since the results of Schmidt (2007) imply that L1 English learners are liable to link English voiced stops with Korean fortis stops, the first prediction is that English voiced stops will show no significant change in VOT over the course of the Korean language program, as they are characterized by short VOTs that are probably not distinguishable from the VOTs of Korean fortis stops. On the other hand, since L1 English learners most likely link English voiceless stops with Korean aspirated stops, it is predicted that English voiceless stops will undergo rapid lengthening of VOT (due to the immediate nature of phonetic drift, its tendency toward approximation of L2 in adult learners, and the substantially longer VOTs of the L2 Korean aspirated stops). Furthermore, the lengthening of VOT is predicted not to exclusively affect the labial and dorsal voiceless stops, the only English voiceless stops that have Korean counterparts with significantly longer VOTs; rather, it is predicted to affect the English voiceless stops at the level of the natural class, such that voiceless stops at all places of articulation (including the coronal stops) drift to a similar degree.

With regard to the second acoustic dimension examined, f_0 onset, English voiced and voiceless stops are both predicted to rapidly increase in f_0, but there are two possible ways in which this increase may occur. The first possibility is that the f_0 increase will occur via category-to-category linkages between English voiced stops and Korean fortis stops, and between English voiceless stops and Korean aspirated stops. Here the English voiced and voiceless stops are both predicted to drift upwards in f_0, for the same reasons that the voiceless stops are predicted to lengthen in VOT: the immediate nature of phonetic drift, its tendency toward approximation of L2 in adult learners, and a noticeably dissimilar L2 norm (a substantially higher f_0 for the Korean fortis and aspirated stops). In this case, the f_0 increase should be limited to English words beginning with voiced and voiceless stops (as well as other segments that could be linked to laryngeally marked Korean consonants, such as /h/); as such, it is not expected to extend to vowel-initial English words. On the other hand, a second possibility is that the f_0 increase will occur at a global level, similar to the way in which drift in the formant frequencies of L1 vowels was hypothesized to occur in Chapter 2. In this case, the f_0 increase would not be limited to English words beginning with voiced and voiceless stops, but instead would extend to all English words including vowel-initial ones. Details about the properties of f_0 in English and Korean specifically do not allow us to adjudicate between these two possibilities; however, the findings of the speech adaptation studies discussed in Chapter 2 (Section 2.2.1), which indicate that f_0 is modulated at least in part by a control mechanism separate from the internal model of segmental control, suggest that changes in the production of f_0, rather than being tied to properties of specific L2 segmental categories, may occur more generally. Thus, the prediction of this study is that phonetic drift in f_0 will occur at a global level, such that f_0 increases not only in stop-initial English words, but also in vowel-initial English words.

In short, phonetic drift in the production of English stop consonants is predicted to occur in a rapid and assimilatory fashion, in accordance with the hypotheses of Chapter 2. While English voiced stops are not expected to change in VOT, since they are already very similar to the perceptually linked Korean fortis stops in this respect, English voiceless stops (at all places of articulation) are expected to rapidly lengthen in VOT under influence from the longer VOT of the perceptually linked Korean aspirated stops. English voiced and voiceless stops are both expected to increase in f_0 onset under influence from the higher overall f_0 level of Korean, which is expected to result in upward f_0 drift in vowel-initial English words as well. It should be

Table 4.2: Predictions for rapid assimilatory phonetic drift in English voiced and voiceless stop consonants, based on the hypotheses presented in Chapter 2. Summarized for both VOT and f_0 onset are: (a) whether drift is predicted, (b) the direction of drift predicted, and (c) the L2 Korean property predicted to trigger the drift.

	VOT			f_0 onset		
Stop type	Drift	Direction	L2 trigger	Drift	Direction	L2 trigger
voiced	no	—	(fortis VOT)	yes	↑	global f_0
voiceless	yes	↑	aspirated VOT	yes	↑	global f_0

noted that although phonetic drift of English is predicted to occur in approximation to Korean, English and Korean categories are nonetheless expected to be produced distinctly rather than merged, due to the general pressure to maintain cross-linguistic contrast. These predictions are summarized in Table 4.2.

4.2 Methods

Recall from Chapter 3 that the production experiments were conducted weekly starting from one week into the language course participants were taking (Section 3.4). Each week, participants read aloud the same set of Korean stimuli and the same set of English stimuli, and their responses were recorded digitally (Sections 3.5–3.6). Participants' recordings were acoustically analyzed in the manner described in Chapter 3 (Sections 3.7.1–3.7.2). Manual measurements of VOT and f_0 onset were taken on learners' productions of the 15 items beginning with plosives (English voiced and voiceless stops; Korean lenis, fortis, and aspirated stops). Raw f_0 measurements were standardized by participant with respect to their f_0 produced over the entire duration of the study. As four tokens were collected of each item, the data presented in Section 4.3 are based on a total of approximately 60 tokens per learner per week (24 of the English items, 36 of the Korean items). Tokens that were anomalous in some way (e.g., pronounced on a yawn, cough, or sigh) were discarded.[1]

In the interest of making valid comparisons, tokens were divided into bins based on three universal phonetic categories of stop voicing (Keating 1984): voicing that begins prior to release (prevoicing), voicing that begins shortly after release (short lag), and voicing that begins following a long delay after release (long lag). All three phonetic voicing types occur in American English generally, and this tripartite nature of voicing production is reflected in the data of the current study, which showed a trimodal distribution of VOT in learners' productions of English stops (Figure 4.3).

[1] Such tokens were few in number, constituting 1.9% of L2 learners' English tokens, 5.0% of their Korean tokens, and 3.7% of native Korean speakers' Korean tokens.

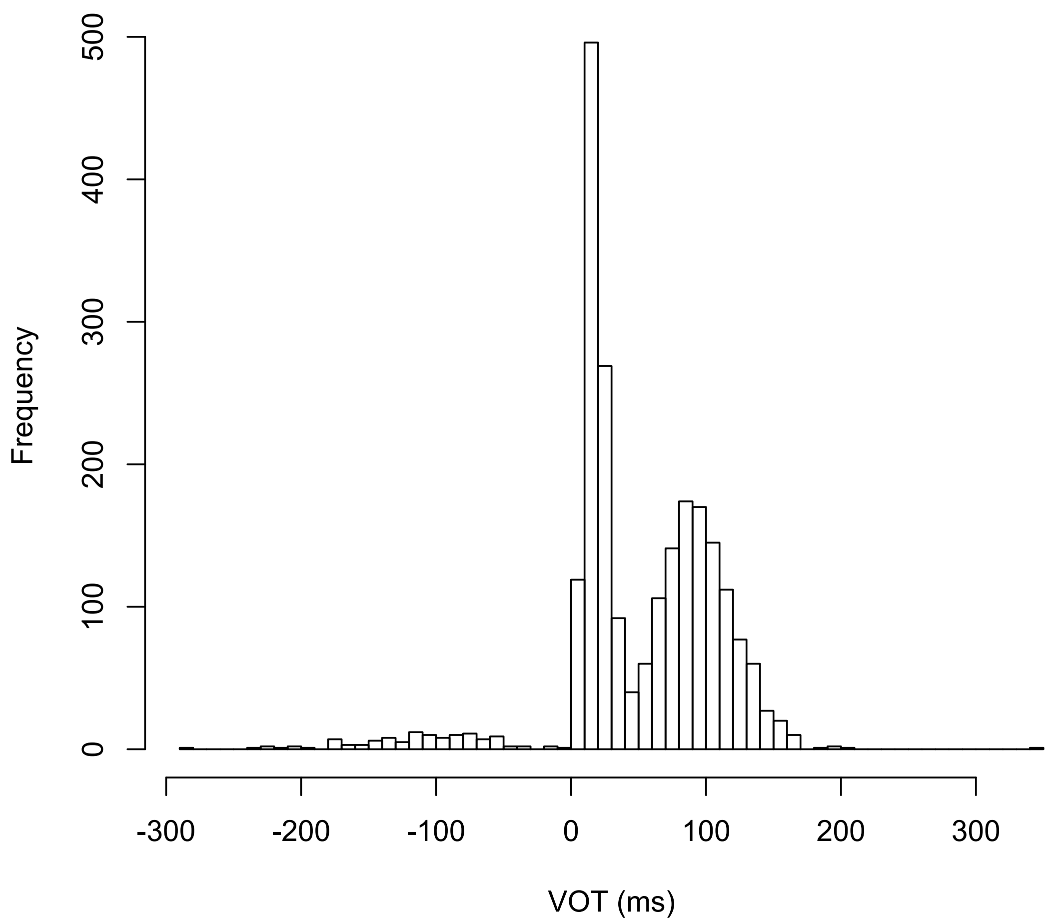


Figure 4.3: Histogram of VOT in learner participants' productions of word-initial English plosives.

Such a distribution was also found in learners' productions of Korean stops, but the distribution of tokens across the three phonetic voicing types was not always equal for perceptually linked English and Korean laryngeal categories. In particular, prevoiced stop tokens, while by far the least common of the three types (occurring at rates between 1% and 13% depending on the laryngeal category and time point in the study), were produced by learners unevenly between the two languages. Consequently, it was not valid to make cross-linguistic comparisons of VOT matching the full set of tokens of one laryngeal category (e.g., English voiced stops, averaged over prevoiced and short-lag tokens) to the full set of tokens of another laryngeal category (e.g., Korean fortis stops, averaged over prevoiced, short-lag, and long-lag tokens), as they did not necessarily have the same distribution of tokens across the three phonetic voicing types.

To circumvent this problem, cross-linguistic comparisons of VOT matched tokens of laryngeal categories by phonetic voicing type (the canonical type for the category, which was the same for paired categories), in accordance with VOT boundaries estimated from the literature (Lisker and Abramson 1964; Lisker, Liberman, Erickson, Dechovitz, and Mandler 1977; Keating 1984) for "prevoiced" stops (< 0 ms), "short-lag" stops (0–30 ms), and "long-lag" stops (> 30 ms). Thus, in the comparison of English voiced stops and Korean fortis stops, which are both typically produced with short-lag VOT, the primary comparison was between short-lag productions. In the comparison of English voiceless stops and Korean aspirated stops, which are both typically produced with long-lag VOT, the primary comparison was between long-lag productions. Other comparisons considered each phonetic voicing type separately (e.g., prevoiced productions of English voiced stops, short-lag productions of Korean lenis stops), rather than averaging over different phonetic voicing types. In every case, it was confirmed that the distribution of data within a phonetic voicing type was normal, and that the quantity of data within a phonetic voicing type stayed relatively stable across time points. Thus, the division of the data into three phonetic voicing types allows for a fairer comparison of the English and Korean categories than can be achieved by averaging over all phonetic voicing types, and in the discussion that follows, it should be kept in mind that—unless otherwise indicated—only canonical phonetic voicing types for the laryngeal categories are being considered: short-lag productions of English voiced stops and Korean fortis stops, and long-lag productions of English voiceless stops and Korean aspirated stops.

4.3 Results

4.3.1 Change in English Voiced Stops

An examination of English voiced stops over time reveals that while VOT does not change substantially, f_0 onset rises significantly (Figure 4.4). A repeated-measures

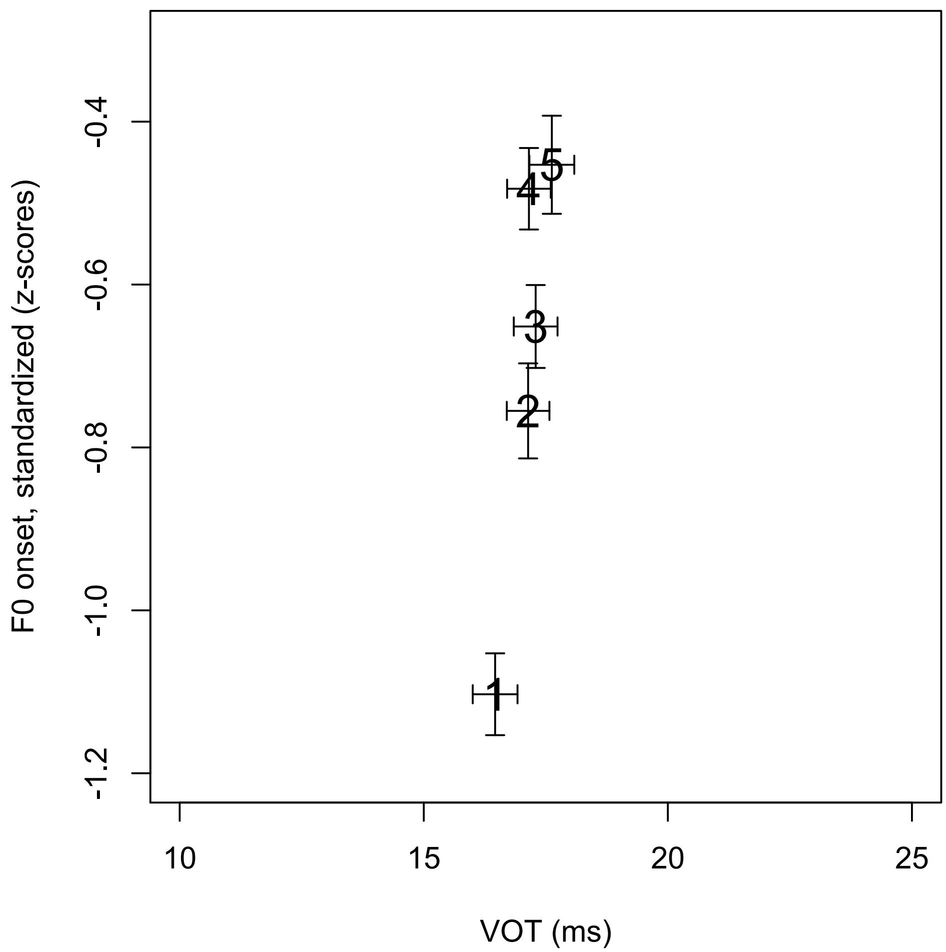
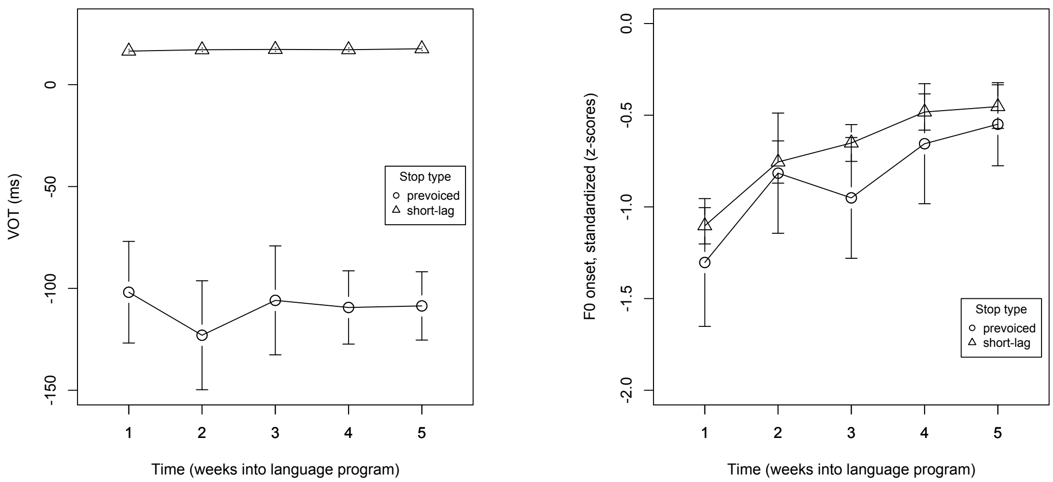


Figure 4.4: Mean f_0 onset by mean VOT in English voiced plosives over time. Numerical symbols plot means of the respective weeks. Error bars indicate ± 1 standard error about the mean.



(a) Mean VOT over time

(b) Mean f_0 onset over time

Figure 4.5: Mean VOT (a) and mean f_0 onset (b) in English voiced plosives over time, by voicing type. Prevoiced productions are plotted in circles, short-lag productions in triangles. Error bars indicate 95% confidence intervals.

analysis of variance (ANOVA) with within-subjects factors Place (of articulation of the stop) and Time (point in the language program) shows a highly significant main effect on VOT of Place [$F(2, 24) = 64.60, p < 0.001$], as expected (Lisker and Abramson 1967; Nearey and Rochet 1994). However, there is no effect of Time [$F(4, 64) = 1.62, n.s.$] and no interaction between these factors [$F(8, 131) = 0.37, n.s.$]. In contrast, a repeated-measures ANOVA reveals a significant main effect of Time on f_0 onset [$F(4, 64) = 5.09, p < 0.01$]. There is a marginally significant main effect of Place [$F(2, 24) = 3.16, p < 0.1$], but no Place x Time interaction [$F(8, 131) = 0.97, n.s.$]. In short, English voiced stops do not increase significantly in VOT, but do increase significantly in f_0 onset, and these effects are not found to differ across the three places of articulation.

It should be noted that the patterns of change in the production of English voiced stops hold not only of short-lag productions, but of prevoiced productions as well. With the exception of a dip in the VOT of prevoiced productions in Week 2, both sets of English voiced stop productions show little change in VOT over time (Figure 4.5a). Moreover, both sets of productions show the same developments in f_0 onset, steadily increasing in f_0 over time (Figure 4.5b). The main effect of Time on f_0 in short-lag productions reaches significance [$p < 0.01$], while the main effect of Time on f_0 in prevoiced productions is marginally significant [$F(4, 10) = 2.61, p < 0.1$]. Post-hoc comparisons of adjacent time points using Tukey's HSD test indicate that for

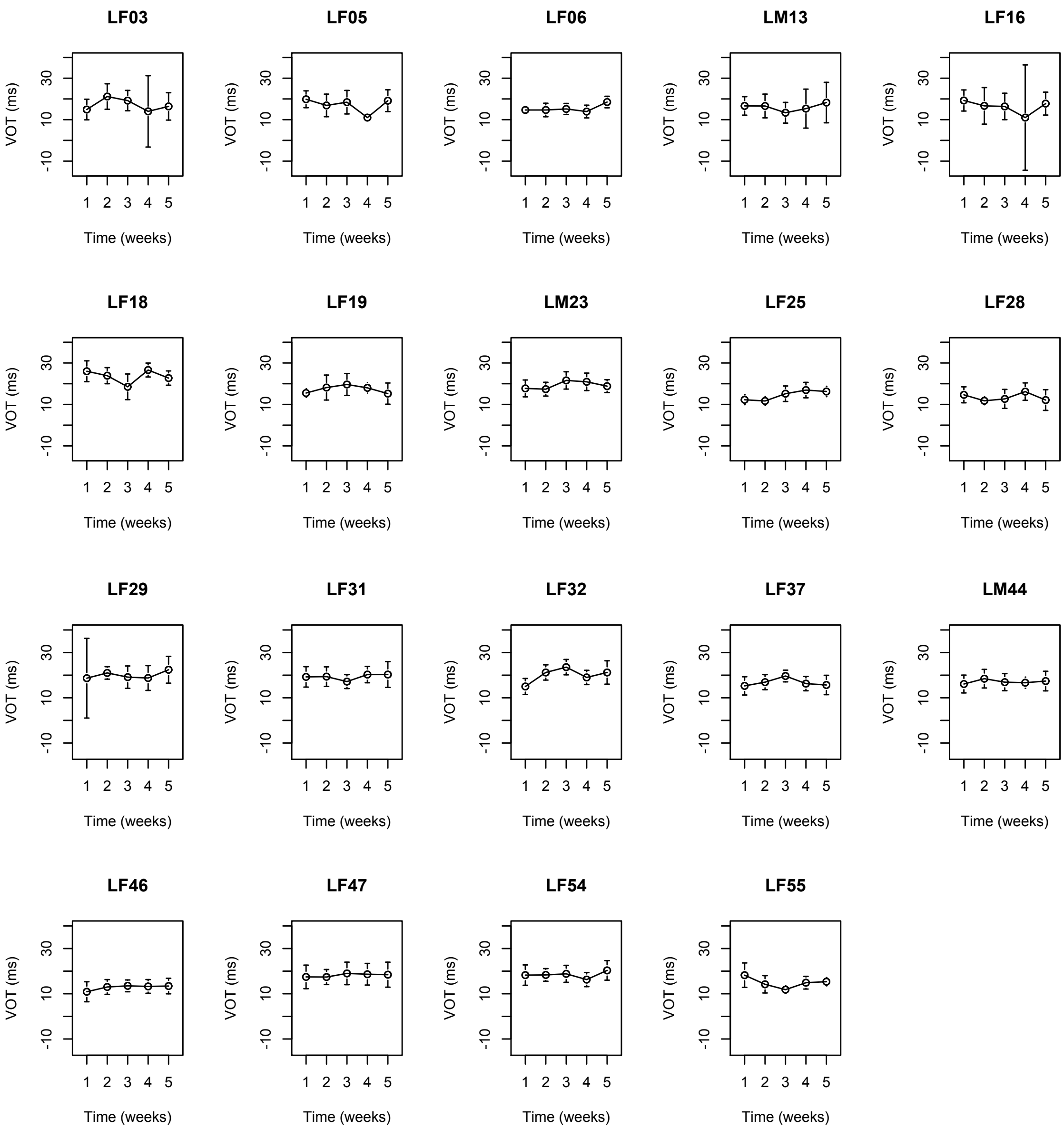


Figure 4.6: Mean VOT in English voiced plosives over time, by participant. Error bars indicate 95% confidence intervals.

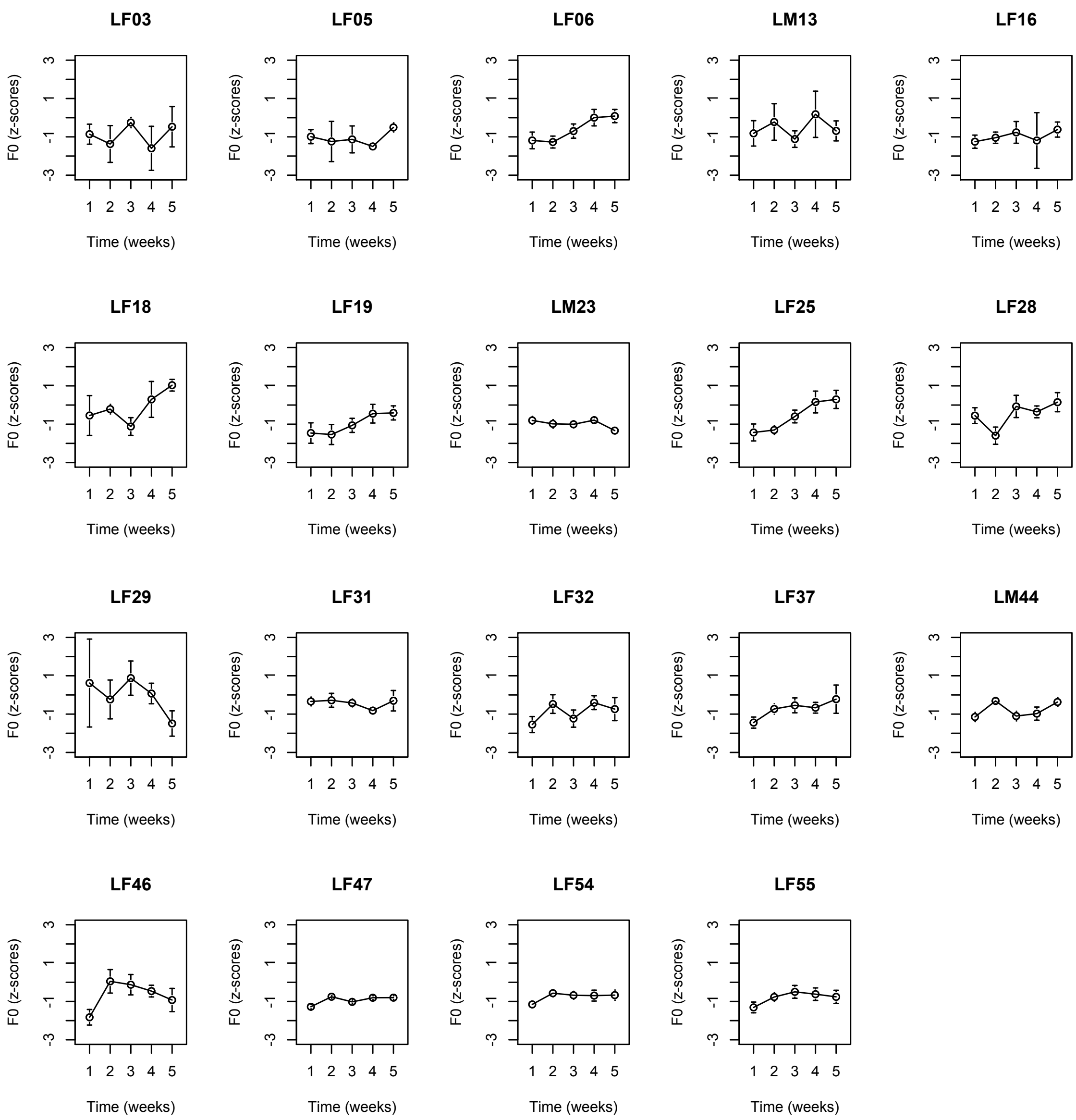


Figure 4.7: Mean f_0 onset in English voiced plosives over time, by participant. Error bars indicate 95% confidence intervals.

short-lag productions, the f_0 changes between Weeks 1 and 2 [$p < 0.001$] and Weeks 3 and 4 [$p < 0.05$] are significant, while for prevoiced productions, the f_0 changes reach significance only between Weeks 1 and 4 [$p < 0.05$] and between Weeks 1 and 5 [$p < 0.01$]. The mean f_0 of prevoiced productions remains below that of short-lag productions in all weeks, as expected.

Finally, the group patterns are generally consistent with those shown by individual participants. In keeping with the lack of significant VOT developments seen in Figure 4.4, few learners show a clear increase or decrease in the VOT of their English voiced stops (Figure 4.6). With the possible exceptions of LF06, LF25, and LF32, who each show a slight increase in VOT, most learners show a pattern of VOT that stays relatively steady over time or fluctuates around a central tendency. Individual developments in f_0 onset are less consistent. However, here too the majority of participants show a pattern resembling the group trend of increasing f_0 seen in Figure 4.4. While two participants (LM23, LF29) actually show the opposite trend, thirteen participants (LF05, LF06, LF16, LF18, LF19, LF25, LF28, LF32, LF37, LF46, LF47, LF54, LF55) show a slight or substantial increase in f_0 onset, with the remaining four participants (LF03, LM13, LF31, LM44) fluctuating around a central tendency (Figure 4.7).

In sum, over the course of Korean classes learners' English voiced stops do not increase significantly in VOT, but do increase significantly in f_0 onset, and the group patterns hold of the majority of individual participants.

4.3.2 Change in English Voiceless Stops

An examination of English voiceless stops over time reveals that both VOT and f_0 onset increase significantly (Figure 4.8). A repeated-measures ANOVA with within-subjects factors Place and Time (the same factors as in Section 4.3.1) shows highly significant main effects on VOT of Place [$F(2, 26) = 9.50, p < 0.001$] and Time [$F(4, 64) = 9.10, p < 0.001$], which do not interact with each other [$F(8, 144) = 1.55, n.s.$]. With respect to f_0, a repeated-measures ANOVA reveals no effect of Place [$F(2, 26) = 1.72, n.s.$], but a significant effect of Time [$F(4, 64) = 4.44, p < 0.01$]; again, there is no interaction between Place and Time [$F(8, 144) = 0.98, n.s.$]. Post-hoc comparisons of adjacent time points using Tukey's HSD test indicate that the VOT changes between Weeks 1 and 2 [$p < 0.01$] and Weeks 4 and 5 [$p < 0.05$] are significant, while the f_0 changes between Weeks 1 and 2 [$p < 0.001$] and Weeks 3 and 4 [$p < 0.05$] are significant. In short, English voiceless stops increase significantly in both VOT and f_0 onset, and these effects are neither limited to the first week nor specific to one particular place of articulation.

The ANOVA results suggest that VOT drifts upward in English voiceless stops at all three places of articulation, but do not indicate whether stops at different places of articulation drift by an equal amount. For this reason, post-hoc analyses of VOT differences between Week 1 and Week 5 were conducted using Tukey's HSD test on

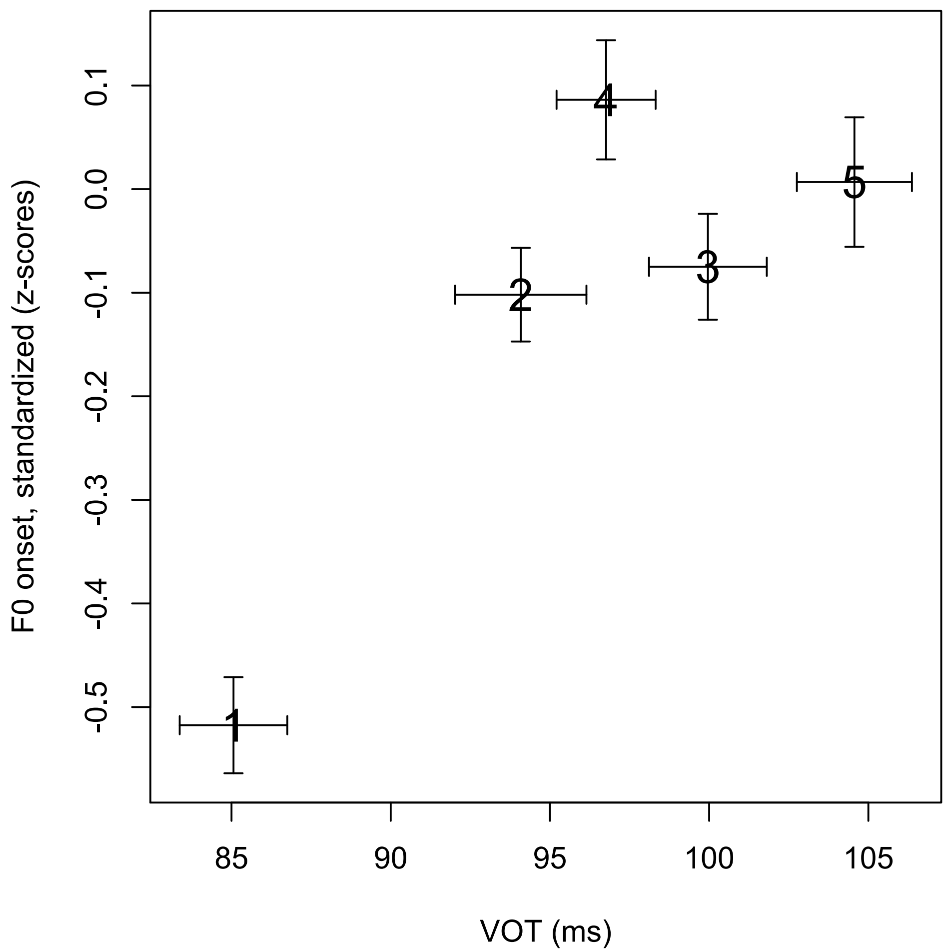


Figure 4.8: Mean f_0 onset by mean VOT in English voiceless plosives over time. Numerical symbols plot means of the respective weeks. Error bars indicate ±1 standard error about the mean.

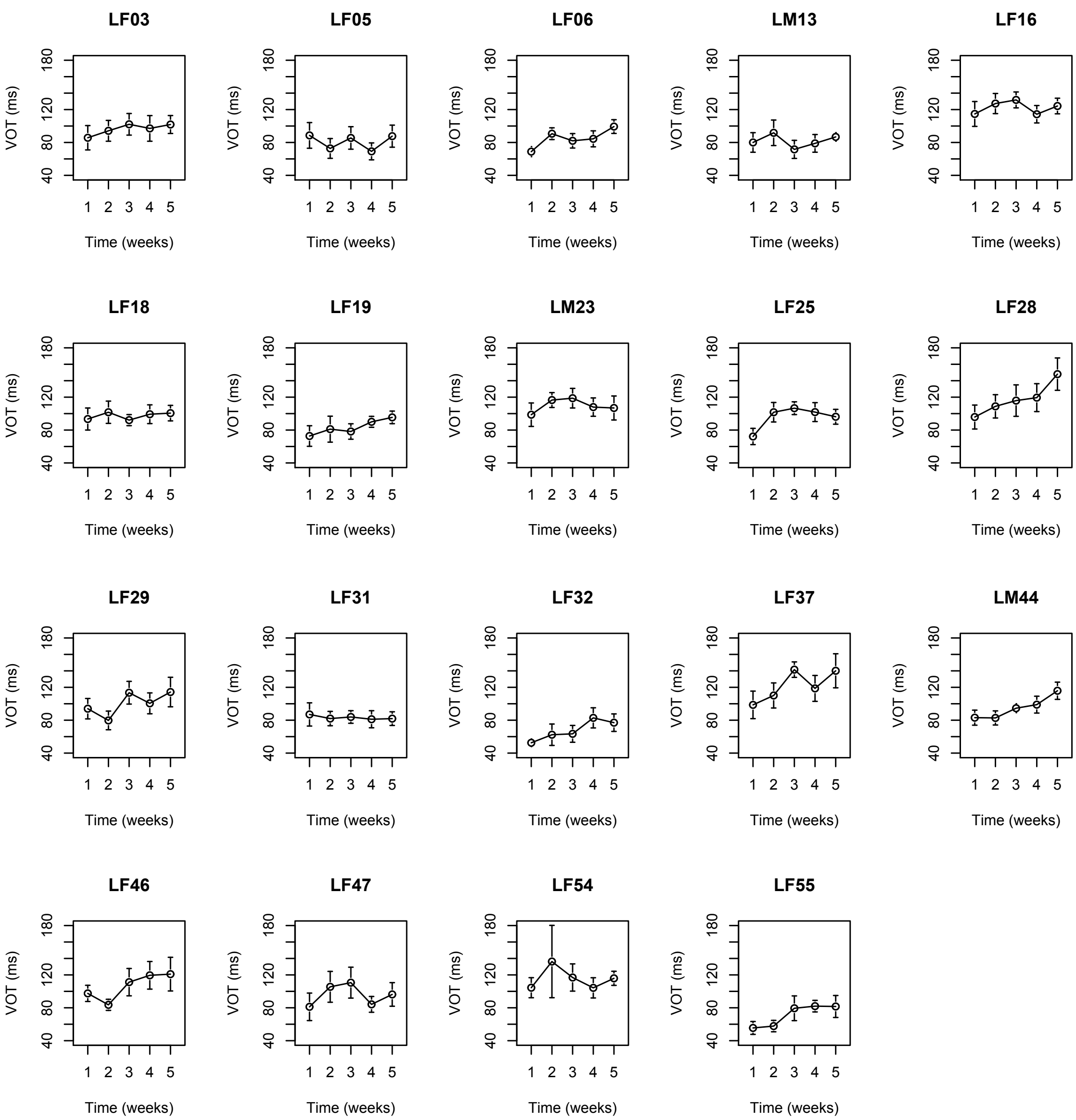


Figure 4.9: Mean VOT in English voiceless plosives over time, by participant. Error bars indicate 95% confidence intervals.

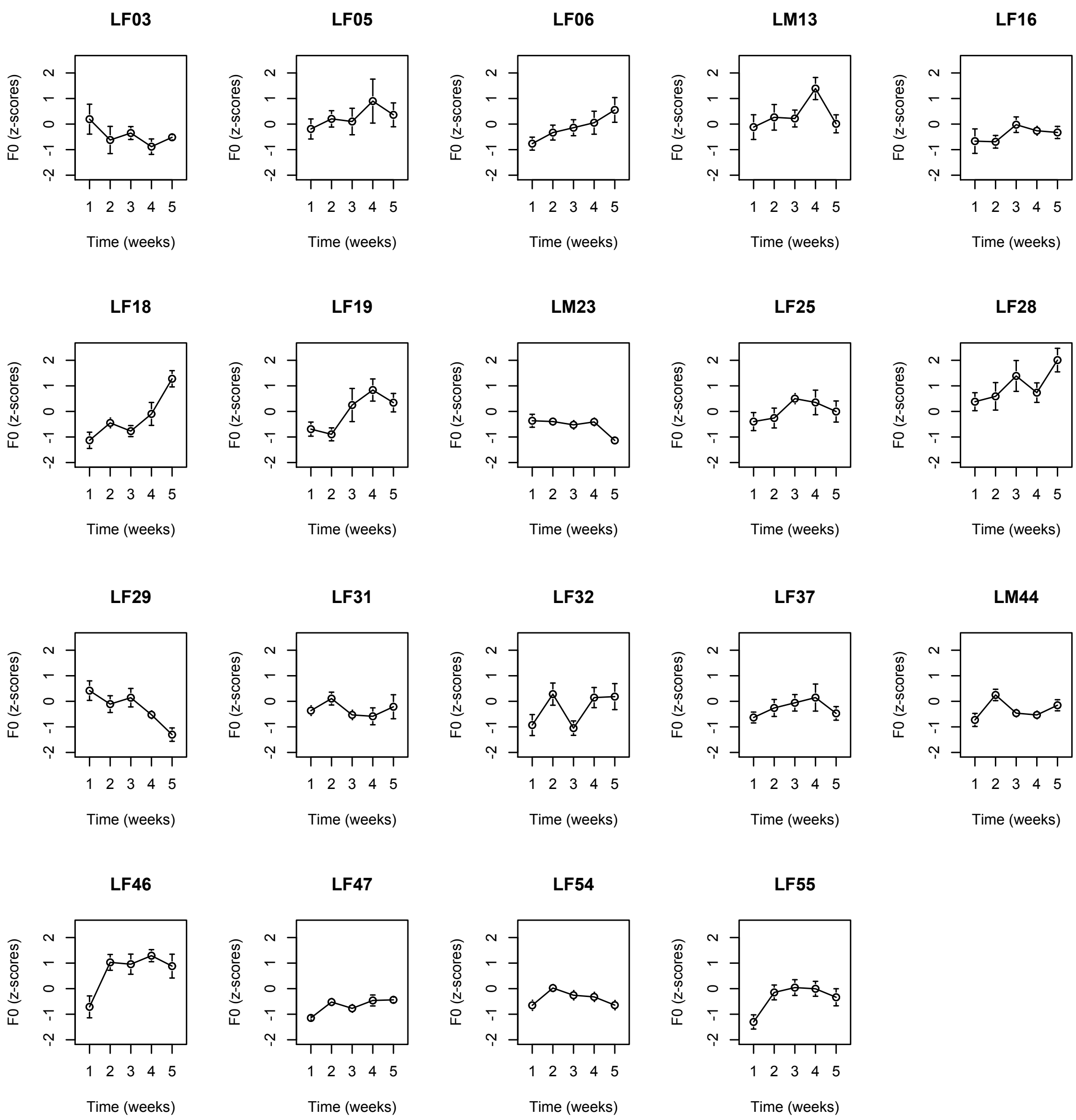


Figure 4.10: Mean f_0 onset in English voiceless plosives over time, by participant. Error bars indicate 95% confidence intervals.

English voiceless stop productions divided by place of articulation. These analyses show that the VOT of English voiceless bilabial stops increases by 22 ms from Week 1 to Week 5, a difference that is highly significant [$p < 0.001$]. The VOT of English voiceless velar stops increases by a similar amount—20 ms—a difference that is also highly significant [$p < 0.001$]. However, the VOT of English voiceless alveolar stops increases by a smaller amount—14 ms—a difference that is only marginally significant [$p < 0.1$]. Thus, these results suggest that while the VOTs of English voiceless stops at all places of articulation lengthen in approximation to the longer VOTs of Korean aspirated stops, the VOT of the English voiceless alveolar stops lengthens to a lesser degree than the VOTs of the voiceless bilabial and voiceless velar stops—a disparity that follows from differences in cross-linguistic similarity. The English voiceless alveolar stops are relatively similar in VOT to the corresponding Korean aspirated denti-alveolar stops, while the English voiceless bilabial and voiceless velar stops are significantly shorter in VOT than the corresponding Korean aspirated bilabial and aspirated velar stops. Consequently, there is less room for drift in the English alveolars, since they have less ground to make up with the Korean denti-alveolars than the English bilabials and velars have with the Korean bilabials and velars.

As with the group patterns for voiced stops, the group patterns for voiceless stops are generally consistent with those shown by individual participants. The group trend of increasing VOT seen in Figure 4.8 holds true of the majority of individual learners (Figure 4.9). Fifteen learners (LF03, LF06, LM13, LF16, LF19, LM23, LF25, LF28, LF29, LF32, LF37, LM44, LF46, LF47, LF55) show an increase in VOT, in comparison to four learners (LF05, LF18, LF31, LF54) whose VOT stays steady or fluctuates around a central tendency. No learners show a decrease in VOT over time. Individual trends in f_0 onset are more variable, but here too the majority of participants show a pattern consistent with the group tendency towards increasing f_0 seen in Figure 4.8. While the opposite trend is found in four participants (LF03, LM23, LF29, LF54), two of whom also show the opposite pattern for voiced stops vis-à-vis the group pattern, thirteen participants (LF05, LF06, LF16, LF18, LF19, LF25, LF28, LF32, LF37, LM44, LF46, LF47, LF55) show a slight or substantial increase in f_0, with the remaining two participants (LM13, LF31) fluctuating around a central tendency (Figure 4.10).

In short, over the course of Korean classes learners' English voiceless stops increase significantly in both VOT and f_0 onset, and these group patterns accurately reflect how the majority of individual participants changed over time. The progression of these phonetic developments relative to changes in learners' L2 Korean production is examined in the next section.

4.3.3 Change in English Compared to Korean

Learners' Korean shows developments over the course of the language program contemporaneous with changes in their English. Comparisons of longitudinal changes

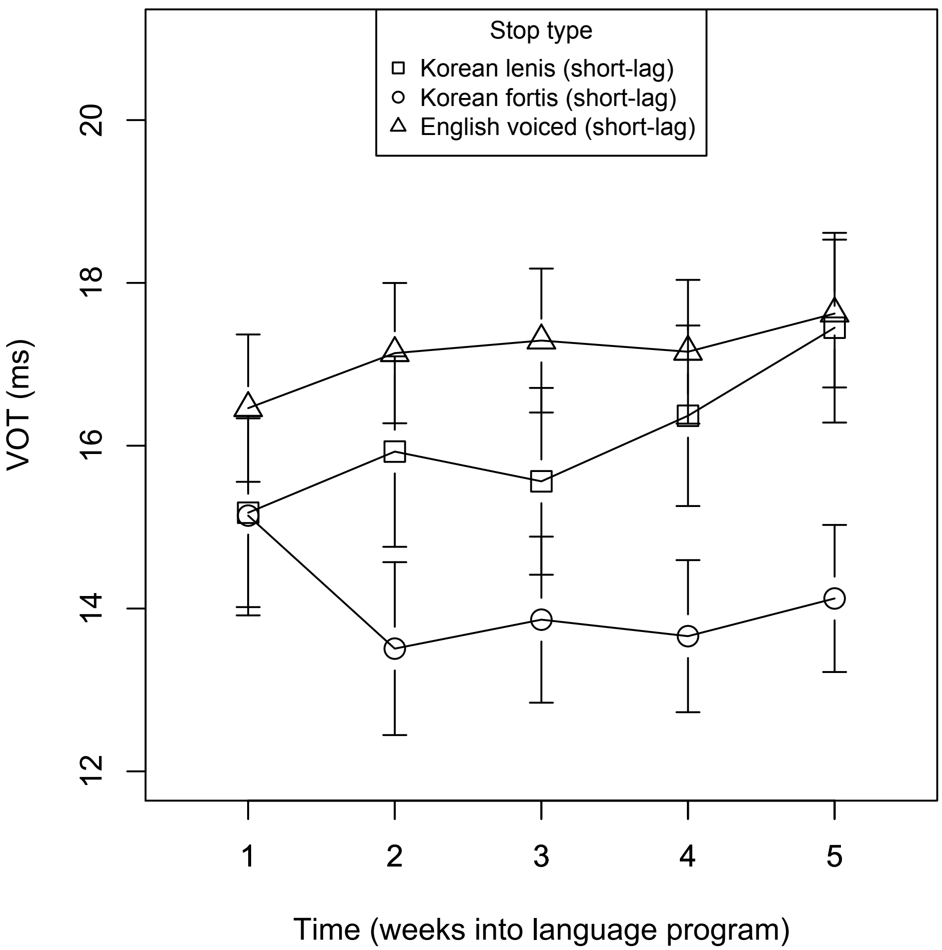


Figure 4.11: Mean VOT in English voiced plosives (triangles), Korean fortis plosives (circles), and short-lag productions of Korean lenis plosives (squares) over time. Error bars indicate 95% confidence intervals.

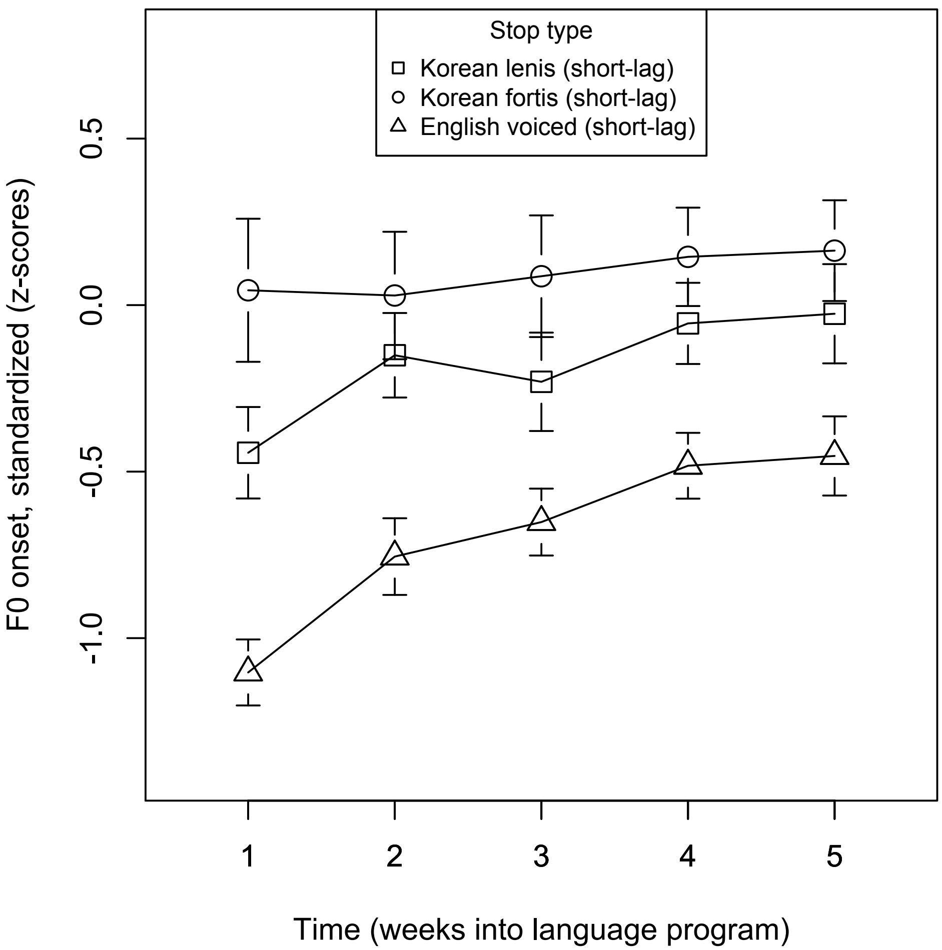


Figure 4.12: Mean f_0 onset in English voiced plosives (triangles), Korean fortis plosives (circles), and short-lag productions of Korean lenis plosives (squares) over time. Error bars indicate 95% confidence intervals.

in L1 English and L2 Korean reveal several noteworthy patterns. Here the main comparisons are between the stop series most likely to be linked perceptually for native English speakers due to being perceived as the most similar in cross-linguistic perception tasks (Schmidt 2007): English voiced stops and Korean fortis stops, and English voiceless stops and Korean aspirated stops. In both cases, the production data are consistent with the predictions presented in Section 4.1.2.

With regard to the short-lag stop series, the VOT of English voiced stops drifts slightly upward while the VOT of Korean fortis stops drifts slightly downward, such that these two stop types, which are not significantly different in VOT in Week 1, are significantly different in VOT by Week 5 (Figure 4.11). The magnitude of the difference between the two categories is small (approximately 4 ms), but significant [$t(352) = 5.74, p < 0.001$ with Bonferroni correction]. Interestingly, the VOT of short-lag productions of Korean lenis stops shows a developmental pattern that is intermediate between that of English voiced and Korean fortis stops. Korean lenis stops start off with the same VOT as Korean fortis stops in Week 1, but then steadily increase in VOT over time to the VOT level of English voiced stops in Week 5. As discussed in Section 4.3.1, there is no main effect of Time on VOT in English voiced stops. Time does not have a main effect on VOT in Korean fortis stops, either [$F(4, 22) = 1.10, n.s.$]. However, Time does have a main effect on VOT in short-lag productions of Korean lenis stops [$F(4, 19) = 3.84, p < 0.05$], attributable to the upward trend seen in Figure 4.11.

As for f_0 onset, the f_0 of Korean fortis stops stays steady, while the f_0 of English voiced stops and the f_0 of short-lag Korean lenis stops both drift upward (Figure 4.12). In addition to the main effect of Time on f_0 of English voiced stops, there is also a marginally significant main effect of Time on f_0 of short-lag Korean lenis stops [$F(4, 19) = 2.47, p < 0.1$]. As was the case with VOT, the f_0 of Korean lenis stops patterns in between that of English voiced and Korean fortis stops, staying at an intermediate level in every week. The result of these f_0 developments is that English voiced stops, along with short-lag Korean lenis stops, become more similar to Korean fortis stops in f_0. The standard f_0 distance between English voiced and Korean fortis stops starts off at 1.2 standard deviations in Week 1, but shrinks to 0.6 standard deviations by Week 5.

While there are only minute changes in the VOTs of English voiced, Korean fortis, and short-lag Korean lenis stops, there are substantial increases in the VOTs of English voiceless and Korean aspirated stops (Figure 4.13). In contrast to the VOT of long-lag productions of Korean lenis stops, which remains relatively steady over time, the VOTs of English voiceless and Korean aspirated stops lengthen significantly. Over five weeks, the VOT of Korean aspirated stops lengthens by approximately 25 ms over its initial level in Week 1, while the VOT of English voiceless stops lengthens by 19 ms. These two sets of stops are not significantly different from each other in VOT at any time point, although they begin to pull apart in Week 4, at which point they are marginally different from each other [$t(346) = -1.70, p < 0.1$].

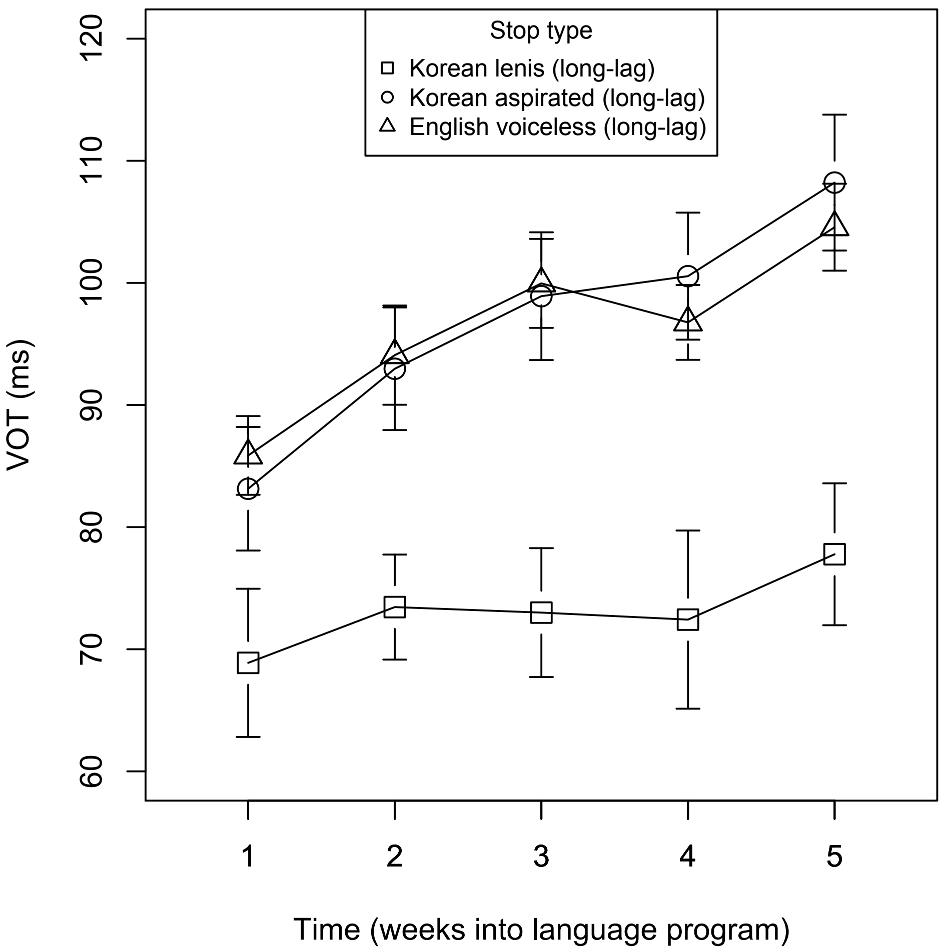


Figure 4.13: Mean VOT in English voiceless plosives (triangles), Korean aspirated plosives (circles), and long-lag productions of Korean lenis plosives (squares) over time. Error bars indicate 95% confidence intervals.

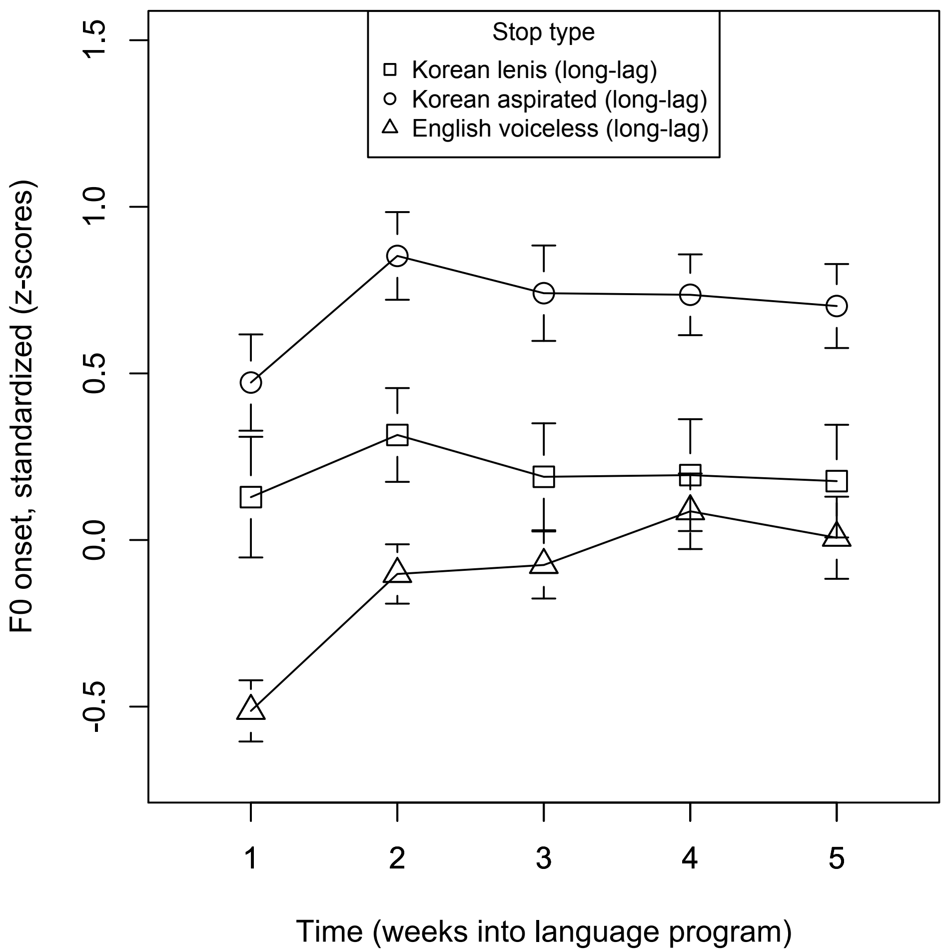


Figure 4.14: Mean f_0 onset in English voiceless plosives (triangles), Korean aspirated plosives (circles), and long-lag productions of Korean lenis plosives (squares) over time. Error bars indicate 95% confidence intervals.

Time has a main effect not only on VOT in English voiceless stops (Section 4.3.2), but also on VOT in Korean aspirated stops [$F(4, 25) = 6.00, p < 0.01$]. This result follows from the pattern seen in Figure 4.13, where it is apparent that the VOTs of English voiceless and Korean aspirated stops increase in a similar way. The initial increase in VOT of Korean aspirated stops is unsurprising, as it is consistent with learners' approximation of the relatively long VOT norm for Korean aspirated stops, which are characterized by VOTs that are over 20 ms longer than those of English voiceless stops on average (Table 4.1). The continued increase in VOT of Korean aspirated stops is more surprising, since by Week 2 learners have already reached native-like VOT levels for this stop series. Thus, by continuing to lengthen VOT beyond this point, learners are actually over-aspirating the Korean aspirated stops. Unlike the increase in VOT of Korean aspirated stops, the increase in VOT of English voiceless stops cannot at any point be explained in terms of phonetic norm approximation, since—at a VOT of 86 ms in Week 1—they start off well above the VOT norm for English voiceless stops, estimated at 69 ms at the beginning of isolated words when averaging over all places of articulation (Lisker and Abramson 1964:391–394). Why do the English voiceless stops increase in VOT then? The way in which Korean aspirated stops and English voiceless stops move in lockstep suggests that they have been perceptually linked to each other (as was predicted), and that what is happening in terms of VOT development is that as the Korean aspirated stops come to be produced with more native-like and then exaggerated VOT, the English voiceless stops "go along for the ride" and increase in VOT as well, even though from the outset this increase results in the English stops becoming less native-like vis-à-vis the phonetic norms of American English.

With regard to f_0 onset, the f_0 of Korean aspirated stops—with the exception of a spike upwards in Week 2—stays relatively steady, as does the f_0 of long-lag productions of Korean lenis stops, which stay at a lower f_0 level overall (Figure 4.14). There is no main effect of Time on f_0 of Korean aspirated stops [$F(4, 25) = 0.84, n.s.$] or on f_0 of long-lag Korean lenis stops [$F(4, 10) = 0.19, n.s.$]. In contrast, the f_0 of English voiceless stops drifts steadily upward, as discussed in Section 4.3.2. The result of this f_0 increase is again that English stops—the voiceless stops in this case—become more similar to Korean stops in f_0. The standard f_0 distance between English voiceless and Korean aspirated stops starts off at 1.0 standard deviations in Week 1, but shrinks to 0.7 standard deviations by Week 5.

The preceding discussion compared the f_0 values of the English stop series to the f_0 values of the perceptually linked Korean stop series and observed that both English voiced and voiceless stops increased in f_0, approximating (though not merging with) the corresponding Korean stops. It is difficult, however, to conclude that these f_0 increases arose via cross-language linkages between laryngeal categories specifically (i.e., English voiced to Korean fortis, English voiceless to Korean aspirated) since the Korean categories do not differ in terms of their potential effect in this respect. The Korean categories are both higher in f_0 than the corresponding English categories;

therefore, they may have triggered upward drift in the f_0 of the English stops via category-to-category linkages or simply via a global link in overall f_0 level across languages. As discussed in Section 4.1.2, what is required to conclude that the observed f_0 increases in English stops resulted solely from cross-language linkages at the level of the laryngeal category is evidence that f_0 does not similarly increase in English words that should be unaffected by the f_0 properties of Korean stop onsets—namely, onsetless (i.e., vowel-initial) words. On the other hand, if f_0 is found to increase in vowel-initial words as well, this would constitute evidence that the observed f_0 increases in English resulted at least in part from a global linkage of overall f_0 level across languages.

Thus, in addition to stop-initial English words, f_0 onset was also measured in the vowel-initial English word *all* according to the same protocols used for stop-initial words (Chapter 3, Section 3.7.2). Contrary to the hypothesis of f_0 drift via category linkage exclusively, f_0 in English vowel-initial productions is also found to increase over time, and the results of a repeated-measures ANOVA with the within-subjects factor Time show that the effect of Time on f_0 here is significant [$F(4, 72) = 4.64, p < 0.01$]. The pattern of f_0 increase in English vowel-initials is similar to the pattern of f_0 increase in English stop-initials (Figure 4.15). However, the magnitude of the overall increase is smaller. English vowel-initials increase in f_0 by 0.38 standard deviations between Week 1 and Week 5, whereas English voiced and voiceless stops increase by 0.66 and 0.53 standard deviations, respectively. Therefore, these data suggest that while the upward drift in the f_0 of English voiced and voiceless stops was influenced by a general increase in English f_0 level approximating the higher f_0 level of Korean, category-to-category linkages to Korean fortis and aspirated stops played a role as well, resulting in greater f_0 drift in English stop-initial words than in English vowel-initial words.

To summarize, cross-linguistic comparisons of developments in VOT and f_0 in English and Korean stop consonants are consistent with the claim that in the minds of the L2 learners under study, English voiced stops are perceptually linked to Korean fortis stops, and English voiceless stops to Korean aspirated stops. Neither English voiced stops nor Korean fortis stops show a significant change in VOT, although they move slightly away from each other in VOT over time. In contrast, both English voiceless stops and Korean aspirated stops show a significant increase in VOT on the order of 20 ms by the final week of the study. The VOTs of these two stop types show no significant differences from each other over five weeks, although they push apart a bit in the last two weeks of the study. As for f_0 onset, English voiced and voiceless stops both drift upwards in f_0, partly due to an approximation of overall Korean f_0 level that also affects English vowel-initials and partly due to an approximation of the specific f_0 norms of the perceptually linked Korean fortis and aspirated stops. The result of this upward drift in f_0 is that by the end of the study period, the standard f_0 distance between parallel English and Korean stop categories shrinks by 0.3–0.6 standard deviations ($\approx$ 5–10 Hz).

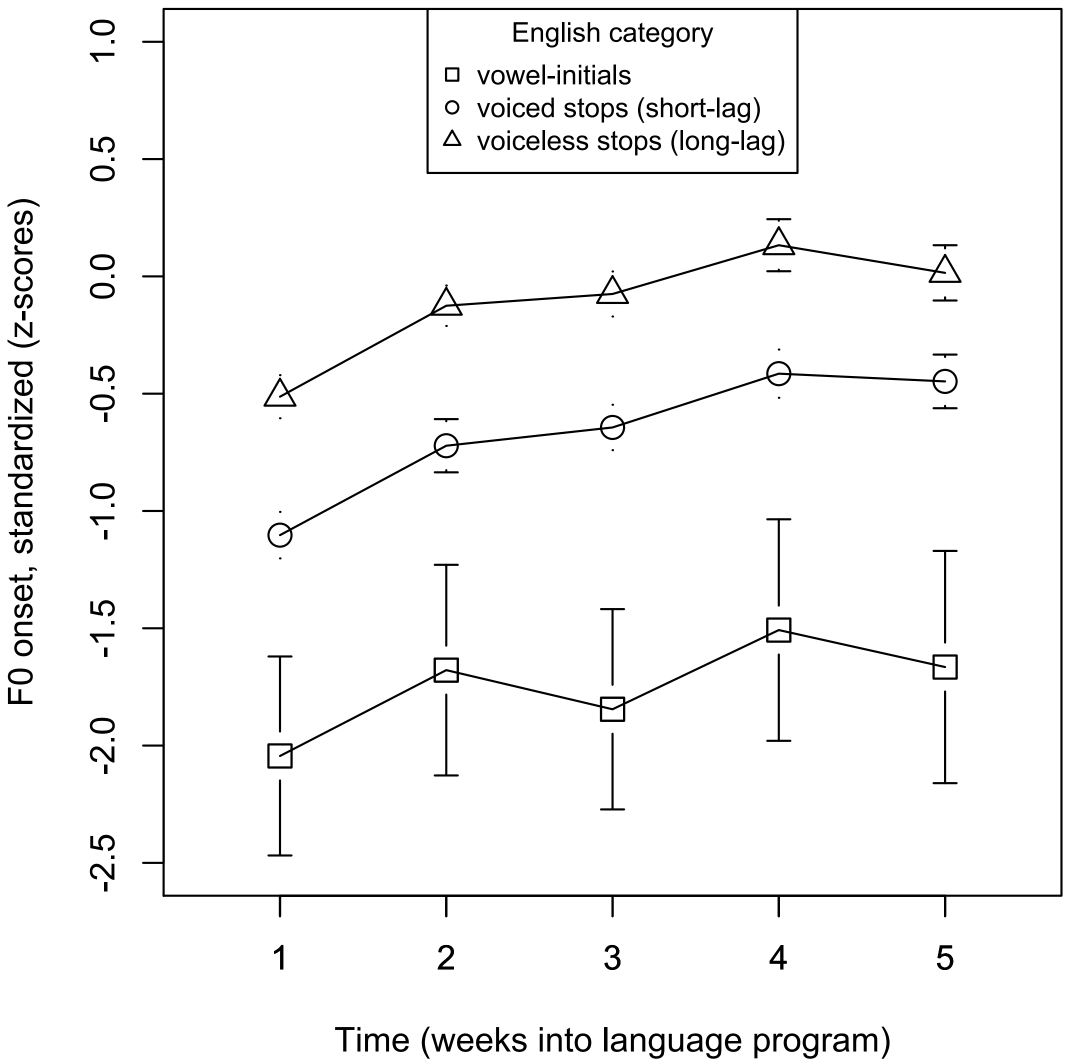


Figure 4.15: Mean f_0 onset in English voiced stops (circles), English voiceless stops (triangles), and English vowel-initials (squares) over time. Error bars indicate 95% confidence intervals.

4.4 Discussion

In this section, the results of this case study are discussed with respect to the general hypotheses of Chapter 2 and the specific predictions of Section 4.1.2. The discussion then moves on to reevaluate the cross-language perceptual linkages postulated in Section 4.1.1, including the role of orthography in establishing these linkages and the cross-linguistic status of Korean lenis stops, before considering some alternative explanations of the findings.

The results of this case study provide evidence supporting the two general hypotheses presented in Chapter 2. First, the results support the hypothesis that L1 phonetic drift, based on automatic equivalence classification and concomitant cross-language linkages, starts early in L2 acquisition. In this case study, significant phonetic drift in learners' English production was found in both VOT and f_0 onset by the second week of the Korean language program. Second, the results support the hypothesis that L1 phonetic drift is assimilatory to L2 for adult learners. As predicted for learners with late-onset L2 exposure, changes in L1 English production approximated the phonetic characteristics of L2 Korean in every case. Despite these patterns of approximation, however, contrast between categories was always maintained, in keeping with the general pressure to maintain cross-linguistic distinctions, and the five English and Korean laryngeal categories in learners' net inventory of stop types were produced distinctly at every time point (Figure 4.16).

The results of this case study also support the four specific predictions presented in Section 4.1.2. First, it was predicted that English voiced stops would not drift significantly in VOT, since their VOT was too similar to that of the perceptually linked Korean fortis stops. This prediction was supported by the data discussed in Section 4.3.1, which showed no significant change in the VOT of English voiced stops. Second, it was predicted that English voiceless stops would rapidly lengthen in VOT, in approximation to the longer VOT norms of the perceptually linked Korean aspirated stops. This result also obtained, with significant increases in the VOT of English voiceless stops being found as early as Week 2 in the Korean language program (Section 4.3.2). Third, it was predicted that drift in the VOT of English voiceless stops would generalize over all places of articulation. Drift did indeed generalize from the bilabial and velar stops to the alveolar stops; however, the alveolar stops still showed less VOT lengthening than the bilabial and velar stops, consistent with the smaller cross-linguistic difference between the VOT norms of alveolars than between the VOT norms of bilabials and velars. Fourth, it was predicted that English voiced and voiceless stops would both drift upwards in f_0 onset as part of a general approximation of overall Korean f_0 level that would also affect vowel-initial English words. In Section 4.3.3, it was shown that English voiced stops, voiceless stops, and vowel-initials all increased in f_0 onset. However, English voiced and voiceless stops increased in f_0 to a greater extent than vowel-initials, suggesting that perceptual linkages to the Korean fortis and aspirated stops also played in role in their f_0 drift. Thus, the picture

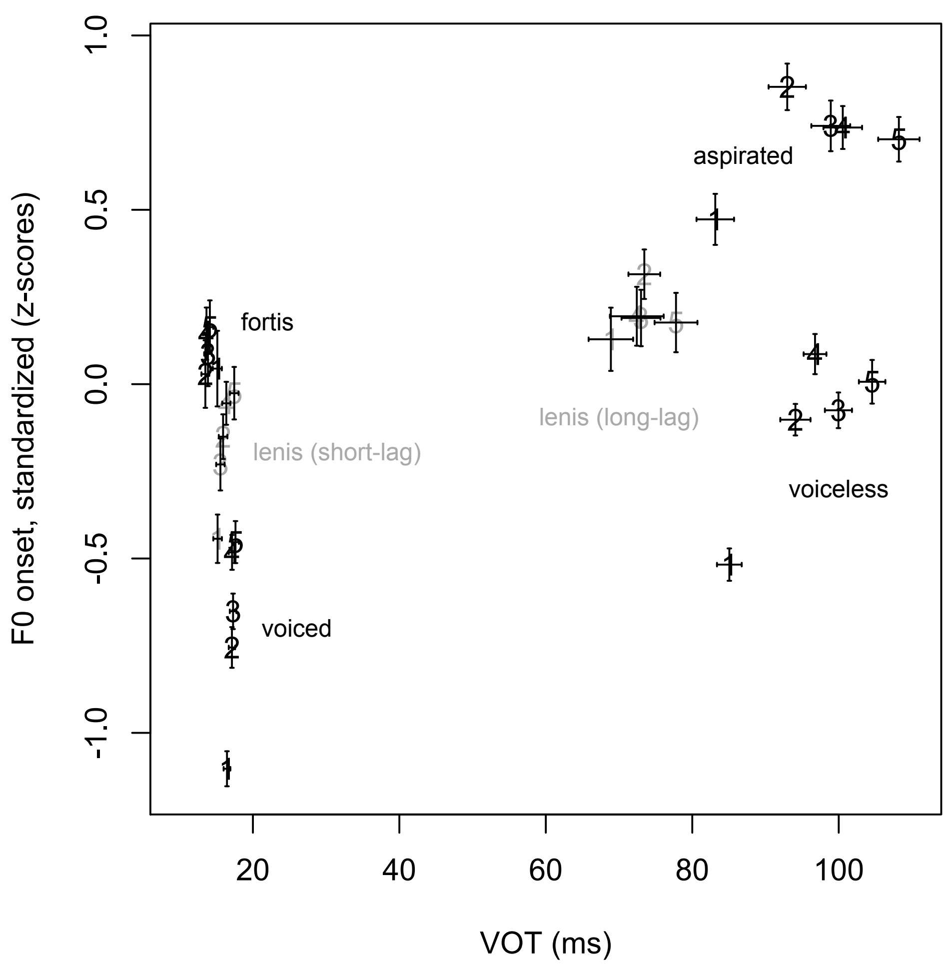


Figure 4.16: Mean f_0 onset by mean VOT in English voiced and voiceless plosives and Korean lenis (short- and long-lag), fortis, and aspirated plosives. Numerical symbols plot the means for the respective weeks. Means for lenis stops are plotted in gray and the rest in black. Error bars indicate ±1 standard error about the mean.

that emerges from these results is one of a multifaceted phonetic drift phenomenon. Contrary to the implications of L2 speech models, which focus almost exclusively on cross-linguistic interference at a segmental level, L2 does not exert influence on L1 via segment-to-segment linkages exclusively; rather, cross-language phonetic effects may arise via linkages at a number of levels—an individual segment, a natural class of several segments, or a global phonetic property.

As discussed in Section 4.1.1, cross-linguistic perceptual data from native English speakers hearing Korean implied that L1 English learners of Korean would perceptually link English voiced stops to Korean fortis stops and English voiceless stops to Korean aspirated stops. This schema of cross-language linkages is consistent with the drift patterns reviewed in Section 4.3.3. It is further supported by evidence from learners' production of Korean fortis stops, which they occasionally pronounce as prevoiced as if they were English voiced stops, suggesting that at some level they think of Korean fortis stops as English voiced stops. This production pattern is unexpected from the point of view of the environment, since prevoiced productions can be assumed to be virtually absent from their Korean input. Korean fortis stops are realized by native Korean speakers as voiceless both word-initially and word-medially, with a relatively long duration of voiceless closure apparent in intervocalic productions (Oh and Johnson 1997). The production of Korean fortis stops like English voiced stops cannot be attributed to orthographic influence either, since it is at odds with the Romanization of Korean fortis stops, as shown in Table 4.3. For the voicing of Korean fortis stops to be based upon an orthographic identification with English voiced stops, Korean fortis stops must be transliterated with graphemes for voiced stops (i.e., <b, d, g>), but they are not. Rather, they are transliterated with graphemes for voiceless stops, as <pp, tt, kk> in the Yale system used by linguists, the McCune-Reischauer system used by non-linguists (Sohn 1999:1–4), and the Revised Romanization system currently used throughout South Korea.[2] Thus, the pattern of voicing Korean fortis stops (which occurs 4–11% of the time depending on the time point in the study) appears to arise ultimately from a perceptually based equivalence classification with English voiced stops, which are prevoiced at similar rates (8–13%) over the course of the study. This suggests that Korean aspirated stops and English voiceless stops, which change in tandem as seen in Figure 4.13, are linked to each other on a similar basis.

Whether Korean lenis stops are linked to a particular English category is less clear, as their phonetic properties, relative to those of English voiced and voiceless stops, are somewhat intermediate. On the one hand, Korean lenis stops are more similar in VOT to English voiceless stops than to English voiced stops in initial position, where they are typically realized with a considerable amount of aspiration (Table 4.1). On the other hand, the relatively low f_0 onset of initial Korean lenis stops

[2] The Revised Romanization instituted by South Korea's Ministry of Culture and Tourism (http://www.korean.go.kr/eng/roman/roman.jsp) is now the official and most widespread system of transliteration in South Korea.

Table 4.3: Transliteration of Korean stop consonants in the Yale, McCune-Reischauer, and Revised Romanization systems.

Consonant	Yale	McCune-Reischauer	Revised
lenis /p͉/	p	p, b	b
lenis /t͉/	t	t, d	d
lenis /k͉/	k	k, g	g
fortis /p͈/	pp	pp	pp
fortis /t͈/	tt	tt	tt
fortis /k͈/	kk	kk	kk
aspirated /pʰ/	ph	p'	p
aspirated /tʰ/	th	t'	t
aspirated /kʰ/	kh	k'	k

makes them similar to English voiced stops. Furthermore, in medial position Korean lenis stops are usually more similar in VOT to English voiced stops, with voicing during closure or a short-lag VOT (Silva 1992, among others). It could be that—with Korean fortis stops linked to English voiced stops and Korean aspirated stops linked to English voiceless stops—Korean lenis stops are the odd category out, constituting a "new" L2 category in the framework of the Speech Learning Model (Flege 1995). However, the findings of Schmidt (2007), which show that English speakers perceive Korean lenis stops as more similar to English voiceless stops than to English voiced stops, but as less close to English voiceless stops than Korean aspirated stops are perceived as being, suggest that, to L1 English learners of Korean, Korean lenis stops more likely constitute a marked version of English voiceless stops than a "new" stop type—in the framework of the Perceptual Assimilation Model (Best 1995), assimilable to an L1 category via Category-Goodness Difference type assimilation, rather than simply being Non-Assimilable. Nevertheless, it is not clear that Korean lenis stops are linked to English voiceless stops, at least in the same way that Korean aspirated stops seem to be. By showing only modest changes in VOT and f_0 limited to short-lag productions, Korean lenis stops in this study do not pattern quite like English voiceless stops (which show substantial changes in both VOT and f_0) or like English voiced stops (which show a significant change in f_0, but not VOT). Thus, it appears that Korean lenis stops either escape linkage with an English category or are so inconsistently classified by learners that no clear group pattern emerges. Future work should attempt to distinguish between these two possibilities.

As the first study to document early phonetic drift in adult L2 learners, the current study has obtained results that are remarkable in that they cannot be argued to be a case of L1 attrition stemming from lack of use, as many previous findings of

L1 phonetic drift could be argued to be. Learners report that although they were learning Korean intensively over the duration of the language program, they spoke English the majority of the time. First-hand observations of participants' activities in and out of class are consistent with these reports, suggesting that the L2 acquisition situation in which learners found themselves is most accurately described as a cross between Second Language Acquisition and Foreign Language Acquisition, as discussed in Chapter 3 (Section 3.3). The fact that participants continued to function predominantly in English while taking Korean classes is to be expected, since even by the end of the program, they had gained only a rudimentary command of the language. However, in spite of the low L2 level attained, L1 phonetic drift still occurred, even in the first week of classes when learners knew very few lexical items of the L2. This suggests that, contrary to what has been assumed in the L2 speech literature, a high level of L2 proficiency is not a prerequisite for L1 phonetic drift.

Though these findings are attributed here to influence from L2 via cross-language links at multiple levels (segmental level, natural-class level, global level), three alternative explanations for the results should be addressed. The first alternative account is that the changes in L1 arose not from relations to the developing L2, but from the experience of having to communicate with Korean locals, who are usually L2 English speakers.[3] In other words, could the L1 developments be attributable to the increased usage, and thus generally higher activation level, of "foreigner talk" (Ferguson 1975)? While the changes in the English voiceless stops can be explained this way, the changes in the English voiced stops cannot, since if the changes merely arose from hyperarticulated pronunciations, one would expect the English voiced stops to shift in ways opposite to the attested patterns: English voiced stops should become more voiced (i.e., at least prevoiced stop productions should decrease in VOT), as well as lower in f_0 onset. However, what actually happened is that English voiced stops changed little in VOT (in fact, the VOT of short-lag productions slightly *increased*) and rose in f_0 onset. Thus, the present results are not likely to have arisen due to the influence of foreigner-directed speech forms.

A related possibility is that, rather than making learners engage in "foreigner talk" (i.e., tend toward clear speech), the experience of communicating with non-native English speakers in Korea resulted in phonetic convergence to accented English. There are two reasons why an account in terms of phonetic accommodation to Korean-accented English is problematic. The first is that phonetic accommodation to non-native talkers appears to be uncommon. As discussed in Chapter 2 (Section 2.2.3), native talkers show the most accommodation to native talkers of their own dialect; they accommodate less to native talkers of a different dialect and are rarely found to accommodate to non-native talkers (Kim 2009). However, even assuming, for the sake of argument, that substantial phonetic accommodation to Korean-accented English talkers would occur in this situation, this account is untenable because it

[3]Thanks to Ann Bradlow for bringing up this possibility.

is inconsistent with the timing of learner participants' interaction with non-native English speakers as such. The first significant interaction of learners with Koreans in English (not Korean) occurred in Week 3 of the language program (when they began teaching English to Korean students), yet production changes are found well before this point. Thus, it is not possible for the results to be attributed to convergence with Korean-accented English.

The third alternative account is that increased familiarity with the experimental task and stimuli—combined with a higher level of comfort with the experimental environment—led to the L1 changes observed by way of allowing participants to give more confident pronunciations.[4] Indeed, higher levels of speaker confidence might very reasonably have resulted in higher levels of f_0 onset; one can even imagine that more confident speakers might have aspirated their voiceless stops more. However, it is unclear why this increased confidence should have affected only English stop productions and not Korean stop productions. As shown in Figures 4.12 and 4.14, with the exception of short-lag productions of Korean lenis stops, Korean stops do not show a clear trend of increasing f_0 onset over time. If anything, though, one would expect Korean stops to most clearly show the effects of increased confidence, as the Korean forms were the ones which were actually unfamiliar to learners to begin with. The fact that they experienced such little change in comparison to English stops suggests that the present results are not the product of gradually increasing levels of speaker confidence.

Finally, the issue of task order should be addressed as well, since the order of Experiments 1K and 2E, in which the Korean experiment always preceded the English experiment, was purposefully kept the same across participants, rather than counterbalanced. This aspect of the experimental design might be a cause for concern because in cross-sectional experiments, the control against effects of particular experimental conditions (e.g., task order) is to vary these conditions across participants so that patterns in the data cannot be attributed to the particular conditions that participants were subject to. On the other hand, in longitudinal studies (which typically track only a few participants or often just one), the control against effects of particular experimental conditions is to keep the conditions the same across time points so that any effect of these particular conditions on the measured variable is automatically parceled out (i.e., if conditions are kept the same, then any change observed over time cannot be attributed to differences in conditions between time points). Given that the present study is both cross-sectional (though the results of only one group, late learners, are reported here) and longitudinal, the task order was kept the same.

The question remains: could pronouncing Korean words in Experiment 1K have affected the way participants pronounced the English words in Experiment 2E? The short answer is yes, but this is not a cause for concern for two reasons. First, while it

[4] Thanks to John Ohala for this suggestion.

is possible that completion of the initial Korean production task affected performance on the subsequent English production task, this effect is controlled for by the fact that the task order was the same at all five time points in this study. Thus, if task order were to affect the results, it could only do so via an interaction with time separate from the effect of accruing L2 experience (e.g., stronger effects of the task order on the measured variable with increasing familiarity with the experiment). The reason that the effect of this interaction would have to be separate from the effect of L2 experience is because the effect of L2 experience is one of the factors being investigated. However, there is no immediately apparent reason to posit such a separate interaction. Second, the way in which task order would have affected performance is precisely the object of study: L2 influence via cross-language linkages. To put it another way, if participants' production of English stops was influenced by their production of Korean stops in a prior task, this result would still be consistent with the arguments advanced in this study, since this short-term influence should also occur via cross-language linkages between L1 and L2. What such an effect of task order would call into question, then, is simply the magnitude of the drift observed. Nevertheless, there is good reason to believe that the preceding Korean production task in fact had no effect on the English production task, and this is discussed further in Chapter 6.

Chapter 5

Phonetic Drift in Vowels

5.1 Introduction

In Chapter 4, it was found that adult L2 learners of Korean manifested phonetic drift in their L1 English stop consonants during the first weeks of learning Korean. In this chapter, learners' production of their L1 English vowels is analyzed in order to examine how the drift found in features associated with L1 obstruents (voice onset time, fundamental frequency onset) compares to drift in features of L1 sonorants—namely, the formant resonances of vowel quality categories. The focus here is on the production of the eleven non-rhotacized American English vowels /i, ɪ, e, ɛ, æ, u, ʊ, o, ɑ, ɔ, ʌ/, in comparison to the seven monophthongal Korean vowels /i, ɛ, u, ɨ, o, ʌ, a/.

5.1.1 Cross-Linguistic Differences

Modern Standard Korean can be said to have seven basic, monophthongal vowel qualities, in addition to a number of diphthongs. Although some researchers posit as many as nine or ten basic vowels (e.g., Lee 1993; Yang 1996a,b; Sohn 1999) and others as few as four (Kim 1968), the number of monophthongal vowels in contemporary Korean as spoken by young adults in Seoul is seven. The basic Korean vowel system of ten monophthongs (/i, y, e, ø, ɛ, u, ɯ, o, ʌ, a/) has shrunk to seven monophthongs due to two changes. The first change is diphthongization of the front rounded vowels /y, ø/ to /wi, we/ (see, e.g., Lee 1993), which are typically realized with a fronted onglide as [ɥi, ɥɛ], respectively. The second change is a phonological merger of the mid front unrounded vowels /e, ɛ/ (Hong 1987; Lee 1995; Ingram and Park 1996, 1997). Although this merger was described as only partial by earlier researchers, recent work by Ko (2009) confirms that the merger is now complete. In addition, there has been a "recent loss of phonemic length" (Kim 2008:42), which has resulted in previously contrastive short and long vowels now being pronounced by most speakers with no reliable difference in duration (Park 1994). Thus, this case study focuses on the

Table 5.1: Native F_1 and F_2 norms for Korean vowels. Figures (in Hz) are means over the seven female and two male native Korean speaker participants. Standard deviations are presented in parentheses alongside the means.

		F_1		F_2	
Vowel	Gender	mean	s.d.	mean	s.d.
/i/	female	337	(28)	2893	(171)
	male	290	(26)	2354	(101)
/ɛ/	female	646	(83)	2409	(135)
	male	482	(43)	1945	(53)
/u/	female	400	(33)	818	(78)
	male	362	(21)	797	(156)
/ɨ/	female	441	(51)	1589	(185)
	male	370	(14)	1401	(51)
/o/	female	445	(46)	778	(249)
	male	402	(34)	699	(39)
/ʌ/	female	686	(84)	1021	(79)
	male	504	(35)	894	(33)
/a/	female	951	(81)	1496	(130)
	male	635	(134)	1106	(80)

potential effect on English vowel production of L2 experience with the seven basic Korean monophthongs /i, ɛ, u, ɨ, o, ʌ, a/.[1] Spectral (F_1 and F_2) norms for these Korean vowels are presented in Table 5.1, averaged over the native Korean speaker participants in the present study.[2]

The American English vowels under consideration comprise the eleven non-rhotacized vowels /i, ɪ, e, ɛ, æ, u, ʊ, o, ɑ, ɔ, ʌ/. The realization of these vowels has been observed to be the locus of much dialectal variation, described in great detail by Labov, Ash, and Boberg (2006:77–116). This dialectal variation is of concern here insofar as it might result in a given Korean vowel being maximally close to

[1]The high back unrounded vowel, often transcribed as /ɯ/ in phonological descriptions of Korean, is transcribed hereafter as /ɨ/ to better represent its central quality in contemporary Seoul Korean.

[2]Yang (1992) also reports acoustic data on native Korean vowels, which differ somewhat from the data in the present study. The main difference is that in the present study, the back vowels are further back in the vowel space. In addition, for male talkers most of the vowels are higher in the vowel space. These discrepancies may be related to the passage of twenty years between the two studies. More likely, however, they are the product of different degrees of L2 experience: Yang's Korean speakers were recorded in the U.S., whereas the Korean speakers in the current study were recorded in Korea. See Chapter 6 (Section 6.4.2) for further discussion.

different English vowels for speakers of different dialects. For this reason, dialectal variation is considered throughout the cross-linguistic comparisons motivating predictions of vowel-to-vowel phonetic drift, with a focus on the dialect regions that have been analyzed in terms of precise, published phonetic norms for vowels (i.e., measures of central tendency for formant values).[3]

Phonetic norms for vowels have not been published on all American English dialect regions, but four acoustic vowel studies have provided data on talkers from the Mid-Atlantic, northern Midwest, South and Southwest, and southern California: Peterson and Barney (1952:183), Hillenbrand, Getty, Clark, and Wheeler (1995:3103), Yang (1996a:250), and Hagiwara (1997:656). Peterson and Barney (1952) did not control for language background strictly, but ended up mostly investigating female talkers who "grew up in the Middle Atlantic speech area" (where the study was conducted) and male talkers who "represented a much broader regional sampling of the United States" and generally spoke "General American" English (Peterson and Barney 1952:177). Hillenbrand et al. (1995), on the other hand, screened many potential participants using a number of methods including questionnaire and production testing in order to focus specifically on talkers who spoke the dialect of the northern Midwest. Most of their speakers "were raised in Michigan's lower peninsula, primarily the southeastern and southwestern parts of the state", and the rest "were primarily from other areas of the upper midwest, such as Illinois, Wisconsin, Minnesota, northern Ohio, and northern Indiana" (Hillenbrand et al. 1995:3099–3100). Yang (1996a) limited his study to talkers from the South and Southwest. No information is provided on these talkers' geographic and dialectal background other than that they "indicated that the American South or Southwest was the area where they spent most of their lives" and "spoke Southern or Southwestern dialects" (Yang 1996a:248); however, it can be assumed that the talkers included in the final sample represented a fairly homogeneous dialect, as potential participants were excluded if their dialect was deemed deviant by a set of same-dialect judges. Moreover, given that they were drawn from students participating in experiments at the University of Texas at Austin, most were probably from Texas. Finally, Hagiwara (1997) examined an ethnically diverse group of college-aged, Southern Californian English speakers with similar geographic, socioeconomic, and educational backgrounds. These speakers represented "a relatively unmarked, middle-class, 'suburban' population" of southern Californians and "as unified a speech community as can reasonably be studied without imposing predetermined sociometric boundaries on a target group of speakers" (Hagiwara 1997:655).

Spectral norms reported in these four studies are summarized in Table 5.2, separated by talker gender. Examination of these formant norms reveals two main differences among them, as summarized by Hagiwara (1997). First, the low front vowel

[3] Though Labov et al. (2006) provide a thorough overview of English vowel variation in the dialects of North America, they discuss the variation in mostly comparative terms, providing ranges for F_1 and F_2, but no measures of central tendency such as means.

/æ/ is produced by talkers from the northern Midwest (NMidW) with a relatively low F_1 and high F_2, resulting in a fronted and raised position in the vowel space relative to its position in the vowel space for talkers from the South/Southwest (S&SW) and southern California (SoCal). The low back vowel /ɑ/ is also produced by NMidW talkers with a high F_2, resulting in a relatively fronted position for this vowel as well. Both of these shifts are features of the Northern Cities Shift characteristic of the NMidW area (Labov 1994). Second, the high back vowels /u, ʊ/, along with the mid central vowel /ʌ/, are produced by SoCal talkers with a high F_2, resulting in fronted positions for these vowels consistent with the California Vowel Shift (Hinton, Moonwomon, Bremner, Luthin, van Clay, Lerner, and Corcoran 1987; Luthin 1987). The high back vowels are produced with a high F_2 by S&SW talkers as well. Consequently, in contrast to the familiar trapezoidal vowel space displayed by the Mid-Atlantic (MidA) talkers in Peterson and Barney (1952), the English vowel space has a triangular configuration for NMidW talkers, while it is shaped like a parallelogram for S&SW talkers and SoCal talkers (Figures 5.1–5.2).

When the English vowel spaces of these four studies are compared to the Korean vowel space, there are both similarities and differences evident in the organization of the acoustic vowel space in the two languages. Korean vowel norms are plotted in comparison to the English vowel norms of the aforementioned English studies in Figures 5.1 and 5.2 for female and male talkers, respectively. In these figures, all of the Korean vowels, as well as the peripheral English vowels that the four English studies have in common, are connected by lines, which highlight the differences in shape between the trapezoidal, triangular, and parallelogram-like English vowel spaces and the heart-shaped Korean vowel space.

With regard to the English high and mid tense vowels, the back rounded /o, u/ have counterparts in Korean /o, u/, but the Korean vowels are consistently found to have a lower F_2 than the English vowels. The F_2 of Korean /o, u/ is lower than that of English /o, u/ for MidA, NMidW, S&SW, and SoCal talkers of both genders, to a degree that English /o/ never falls closest to Korean /o/. Instead, English /o/ generally falls closest to Korean /ʌ/ for NMidW and S&SW talkers and closest to Korean /ɨ/ for SoCal talkers. Meanwhile, the dialectal divide with respect to fronting of English /u/ results in English /u/ generally falling closest to Korean /u/ for MidA and NMidW talkers, but closest to Korean /ɨ/ for S&SW and SoCal talkers. Both languages contain the high front vowel /i/, which is realized similarly in each. English /e/ is located in between Korean /i/ and Korean /ɛ/—closer to Korean /ɛ/ for NMidW, S&SW talkers, and male SoCal talkers, but closer to Korean /i/ for female SoCal talkers.

As for the English high and mid lax vowels, high front /ɪ/ is located in between Korean /i/ and Korean /ɛ/, but closer to Korean /ɛ/ except in the case of female MidA talkers. English /ɛ/ is close to Korean /ɛ/, although lower and more retracted for female NMidW and SoCal talkers and more retracted for male SoCal talkers. High back /ʊ/ is generally realized closest to Korean /ɨ/, although for the male talkers in

Table 5.2: Native F_1 and F_2 norms for American English vowels. Figures (in Hz) are averages over talkers from the Mid-Atlantic (MidA), northern Midwest (NMidW), South and Southwest (S&SW), and southern California (SoCal).

		MidA		NMidW		S&SW		SoCal	
Vowel	Gender	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2
/i/	female	310	2790	437	2761	390	2826	362	2897
	male	270	2290	342	2322	286	2317	291	2338
/ɪ/	female	430	2480	483	2365	466	2373	467	2400
	male	390	1990	427	2034	409	2012	418	1807
/e/	female	—	—	536	2530	521	2536	440	2655
	male	—	—	476	2089	469	2082	403	2059
/ɛ/	female	610	2330	731	2058	631	2244	808	2163
	male	530	1840	580	1799	531	1900	529	1670
/æ/	female	860	2050	669	2349	825	2059	1017	1810
	male	660	1720	588	1952	687	1743	685	1601
/u/	female	370	950	459	1105	417	1511	395	1700
	male	300	870	378	997	333	1393	323	1417
/ʊ/	female	470	1160	519	1225	491	1486	486	1665
	male	440	1020	469	1122	446	1331	441	1366
/o/	female	—	—	555	1035	528	1206	516	1391
	male	—	—	497	910	498	1127	437	1188
/ɑ/	female	850	1220	936	1551	857	1255	997	1390
	male	730	1090	768	1333	694	1121	710	1221
/ɔ/	female	590	920	781	1136	777	1140	—	—
	male	570	840	652	997	663	1026	—	—
/ʌ/	female	760	1400	753	1426	701	1641	847	1753
	male	640	1190	623	1200	592	1331	574	1415

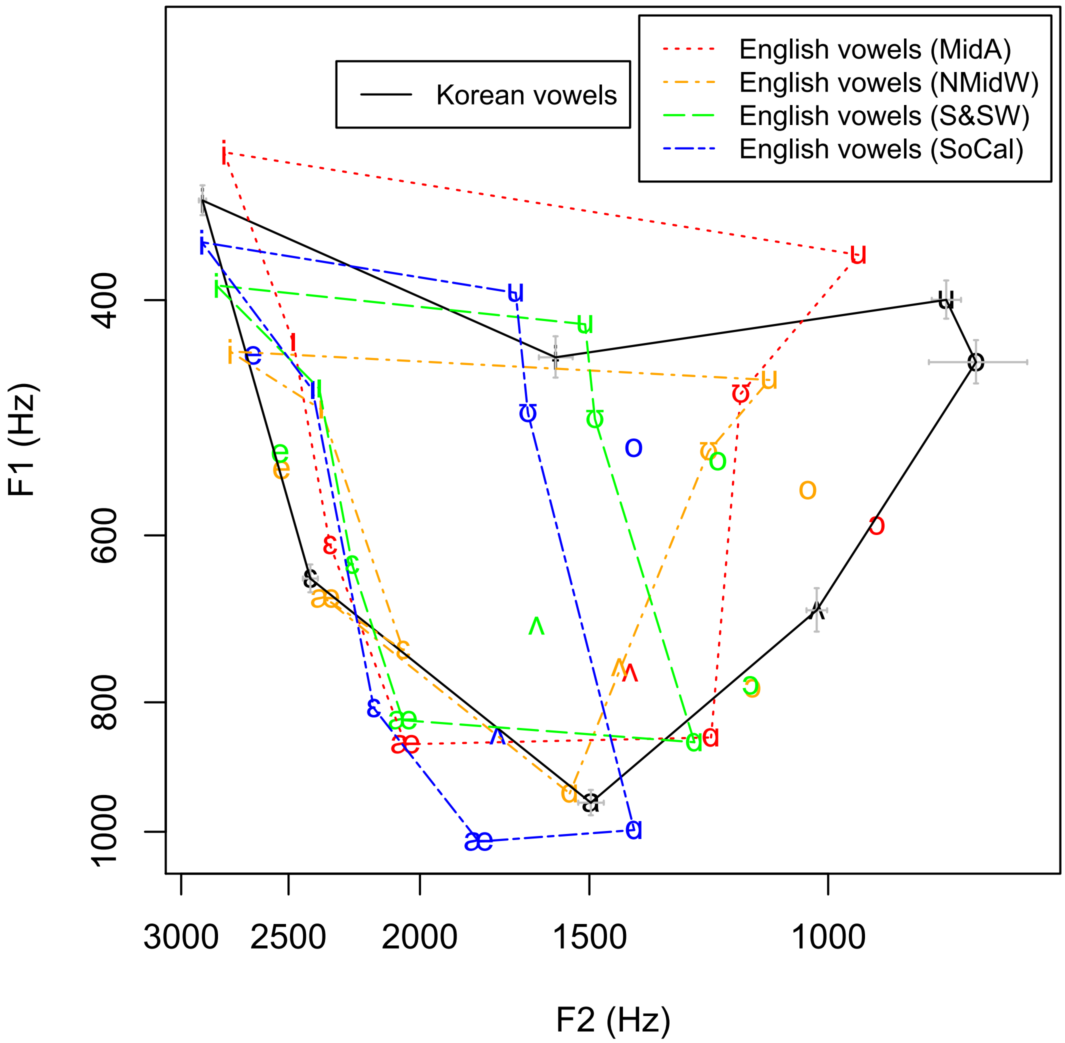


Figure 5.1: Mean F_1 by mean F_2 of Korean vowels and English vowels as produced by female talkers. Korean values are from native Korean speakers in the present study; English vowels are from native English speakers from the Mid-Atlantic (MidA), northern Midwest (NMidW), South and Southwest (S&SW), and southern California (SoCal). The scale of both axes is logarithmic. Plot symbols are the standard IPA transcriptions of the vowels. Error bars indicate ±1 standard error about the mean.

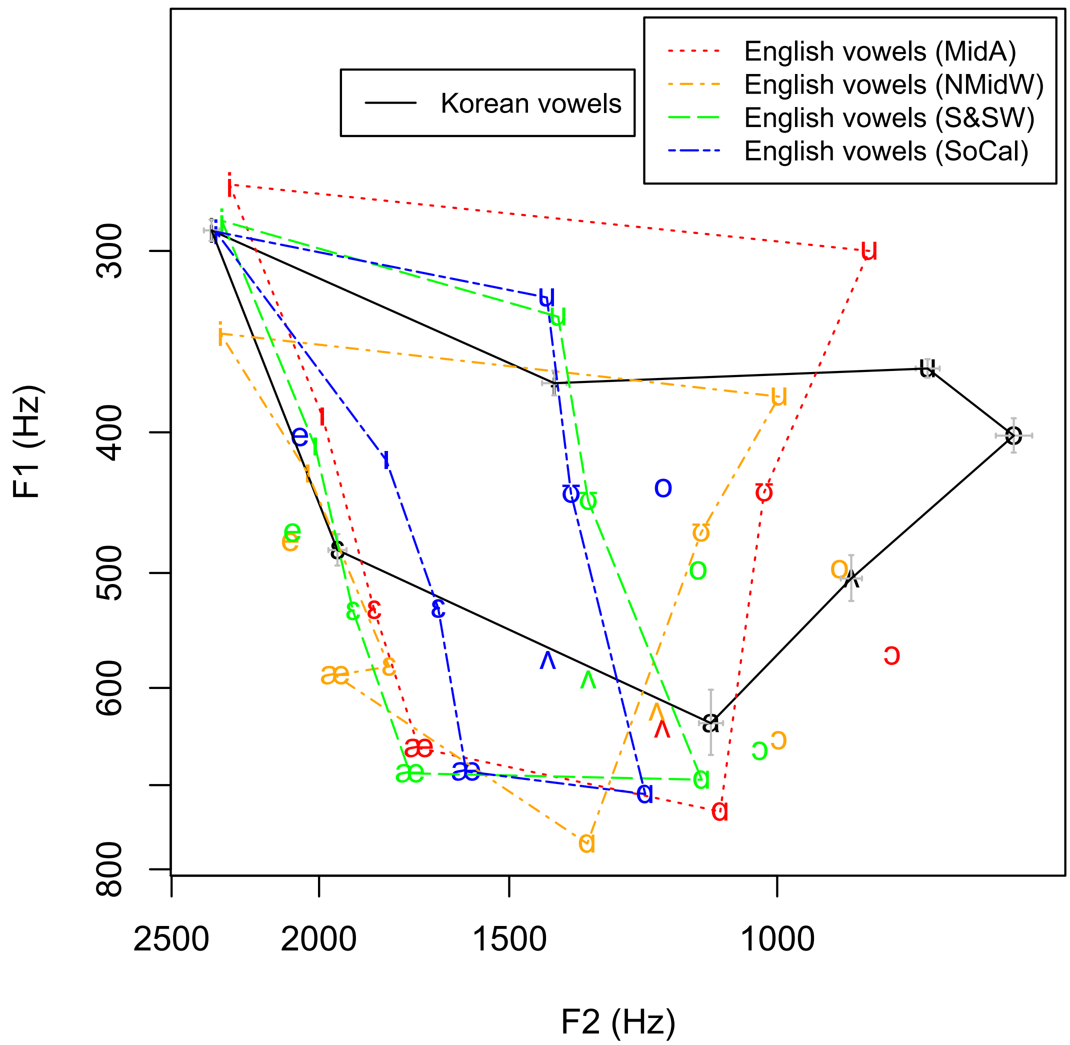


Figure 5.2: Mean F_1 by mean F_2 of Korean vowels and English vowels as produced by male talkers. Korean values are from native Korean speakers in the present study; English vowels are from native English speakers from the Mid-Atlantic (MidA), northern Midwest (NMidW), South and Southwest (S&SW), and southern California (SoCal). The scale of both axes is logarithmic. Plot symbols are the standard IPA transcriptions of the vowels. Error bars indicate ± 1 standard error about the mean.

Peterson and Barney (1952) it is slightly closer to Korean /u/. For S&SW and SoCal talkers, both English /u/ and /ʊ/ are close to Korean /ɨ/, consistent with the fact that "[u] and [ʊ] are unrounded, [ʊ] often being pronounced with spread lips" (Ladefoged 1999:43). Finally, English /ʌ/ lies relatively far from Korean /ʌ/ and is instead closer to Korean /a/. Korean /ʌ/ is generally closest to English /ɔ/, although also relatively close to English /o/; in fact, it is located in nearly the same spot as English /o/ for male NMidW talkers. For SoCal talkers, too, Korean /ʌ/ is closest to English /o/, but it still lies relatively far from both /o/ and /ɑ/, the two closest English vowels.

As for the low vowels, low front /æ/ lies close to Korean /ɛ/, especially for NMidW talkers. English /æ/ is closest to Korean /ɛ/ also for MidA and S&SW talkers, but in the case of SoCal talkers, for whom /æ/ is relatively retracted, /æ/ lies between Korean /ɛ/ and /a/ or closer to /a/. The low back unrounded /ɑ/ of English is produced by female NMidW talkers as quite close to the low central /a/ of Korean; this is also true of female SoCal talkers. For male NMidW talkers, English /ɑ/ is even further front than Korean /a/, a pattern that also holds of male SoCal talkers. For MidA and S&SW talkers, English /ɑ/ is more back, but the closest Korean vowel remains /a/. The low back rounded /ɔ/ is closest to Korean /ʌ/ for MidA talkers, as well as female NMidW and S&SW talkers. On the other hand, for male NMidW and S&SW talkers, English /ɔ/ is closest to Korean /a/.

In sum, despite dialectal variation in acoustic proximity between English and Korean vowels, several cross-language vowel pairs emerge as consistently close across dialects. The closest English-Korean vowel pairs discussed above are summarized in Table 5.3. Relative consistency is expected across dialects with respect to the influence of Korean vowels on English /i, ɪ, e, ɛ, ʊ, ʌ/, as in each of these cases the closest Korean vowel to the English vowel—and, thus, the most likely L2 "attractor"—is, with few exceptions, the same across dialects and in a similar position relative to the English vowel. In the case of English /ɑ/, too, the closest Korean vowel is the same for every dialect—namely, Korean/a/; however, Korean /a/ is positioned differently across dialects: behind the /ɑ/ of NMidW and SoCal talkers, but in front of the /ɑ/ of MidA and S&SW talkers. This discrepancy may result in different patterns of phonetic drift in English /ɑ/ for these two dialect groups. Cross-dialectal variation is also likely to occur in the influence of Korean vowels on English /æ, u, o, ɔ/. In particular, there is a salient disparity in the cross-language proximity of the tense back English vowels /o, u/ due to the discrepancy between S&SW and SoCal talkers' F_2 values and MidA and NMidW talkers' F_2 values.

The preceding survey of cross-linguistic differences between Korean and English vowels focused on differences between the realizations of individual vowels. However, there are also differences between the aggregate vowel systems. English has a greater number of basic vowels than Korean—eleven as opposed to seven. By virtue of this fact, the distribution of vowels within the English vowel space is not the same as the distribution of vowels within the Korean vowel space. To be specific, the English vowel space is more crowded than the Korean vowel space in both the front region

Table 5.3: Cross-language acoustic phonetic proximity between Korean and English vowels. For each English vowel in the speech of talkers from the Mid-Atlantic, northern Midwest, South and Southwest, and southern California, the Korean vowel (or vowels) is given that is the closest acoustically in terms of Euclidean distance in F_1 x F_2 space.

Vowel	Gender	Mid-Atlantic	N. Midwest	South/Southwest	S. California
/i/	female	i	i	i	i
	male	i	i	i	i
/ɪ/	female	i	ɛ	ɛ	ɛ
	male	ɛ	ɛ	ɛ	ɛ
/e/	female	—	ɛ	ɛ	i
	male	—	ɛ	ɛ	ɛ
/ɛ/	female	ɛ	ɛ	ɛ	ɛ
	male	ɛ	ɛ	ɛ	ɛ
/æ/	female	ɛ, a	ɛ	ɛ	a
	male	ɛ	ɛ	ɛ	ɛ, a
/u/	female	u	o, u, ɨ	ɨ	ɨ
	male	u	u	ɨ	ɨ
/ʊ/	female	ɨ	ɨ	ɨ	ɨ
	male	u	ɨ	ɨ	ɨ
/o/	female	—	ʌ	ɨ, ʌ	ɨ
	male	—	ʌ	ʌ, a	ɨ
/ɑ/	female	a	a	a	a
	male	a	a	a	a
/ɔ/	female	ʌ	ʌ	ʌ	—
	male	ʌ	a	a	—
/ʌ/	female	a	a	a	a
	male	a	a	a	a

Table 5.4: Overall F_1 and F_2 levels of the Korean vowel space and the American English vowel space. Figures (in Hz) are averages over norms for the seven Korean monophthongs as produced by the native Korean speaker participants, and for the eleven English monophthongs as produced by talkers from the Mid-Atlantic (MidA), northern Midwest (NMidW), South and Southwest (S&SW), and southern California (SoCal).

		female		male	
Language	Variety	F_1	F_2	F_1	F_2
English	MidA	583	1700	503	1428
	NMidW	624	1776	527	1523
	S&SW	600	1843	510	1580
	SoCal	634	1982	481	1608
Korean	Standard	558	1572	435	1314

and the lower (i.e., non-high) region. With regard to the front region, the English vowel space contains five front vowels, compared to two front vowels in Korean. With regard to the lower region, the English vowel space includes a rather centralized /ɪ/, /e/, /æ/, and, for many speakers, /ɔ/—vowels that are all absent from the Korean inventory. Moreover, several vowels that the two languages have in common (in that they are either standardly transcribed with the same IPA symbol or are phonetically close) are located at different points in the vowel space, with the Korean vowel being realized as higher (and, often, as more back). Korean /o/, /ʌ/, and /ɨ/ are each higher than English /o/, /ʌ/, and /ʊ/, respectively; /o/ and /ʌ/, moreover, are further back in Korean than in English. These comparisons suggest that the Korean vowel space is probably higher and more back overall than the English vowel space, and this conclusion is consistent with overall F_1 and F_2 levels in the two languages (Table 5.4). When the norms for the individual vowels in each system are averaged to calculate grand means for F_1 and F_2 over the entire vowel space, overall F_1 and F_2 levels are consistently found to be lower for the Korean vowel space.[4] In other words, the Korean vowel space as a whole is indeed higher and more back than the English vowel space, and this cross-linguistic difference is true of Korean in comparison to all four of the English dialects examined here.

[4]Norms for English /e, o/ are not reported by Peterson and Barney (1952), though these vowels exist in MidA English. However, the absence of these non-low vowels in the calculation of a grand mean for F_1 in MidA English probably does not affect the average much; if anything, it is likely to lower, rather than raise, the average, thus strengthening the case for lower formant levels in Korean. Norms for the English vowel /ɔ/ are not reported by Hagiwara (1997) since /ɔ/ is generally merged with /ɑ/ in California English.

5.1.2 Predictions

Three hypotheses presented in Chapter 2 are relevant to this case study. The first hypothesis was that L1 phonetic drift would occur early in L2 acquisition because of early-established cross-language linkages based on automatic equivalence classification of novel L2 sounds. The second hypothesis was that L1 phonetic drift would occur in assimilation to L2 due to the late onset of L2 experience in the study participants, who were adult L2 learners. The final hypothesis was that, with respect to vowels specifically, L1 phonetic drift would occur at a global level—that is, at the level of the vowel system, similar to the way in which drift of L1 vowels was found to occur in L1 Quichua-L2 Spanish bilinguals (Guion 2003).

Considering the similarities and differences between the Korean and English vowels described above, these hypotheses lead to two predictions regarding phonetic drift in English vowels. The first prediction is that production of English vowels will generally shift upwards and backwards (i.e., F_1 and F_2 values will be found to decrease) by the end of the Korean language program (due to the immediate nature of phonetic drift, its tendency toward approximation of L2 in adult learners, and the lower overall F_1 and F_2 levels of Korean). If this shift is accomplished via global linkages to overall F_1 and F_2 levels in Korean, the implication is that shifts in the production of English vowels will not be explicable in terms of shifts toward individual Korean vowels that are nearby in the acoustic vowel space, since the Korean vowels closest to English vowels do not all happen to be located in the same position relative to the English vowels, as discussed in Section 5.1.1. The second prediction is that, although the magnitude of phonetic drift in vowel production may differ across dialects, the direction of drift will not differ, since the dialects examined are consistent in having higher overall F_1 and F_2 levels than Korean.

If, on the other hand, phonetic drift in L1 vowel production occurs at the level of individual vowels, then L1 vowels are expected to drift towards L2 vowels that have been perceptually linked to them, but stop short of merging with them (in order to maintain cross-linguistic distinctions). In this case, the L2 vowel that is perceptually linked to an L1 vowel will usually be the closest acoustically (Table 5.3). However, given that perceptual assimilation of vowels has been shown not to follow straightforwardly from acoustic phonetic proximity (Polka and Bohn 1996; Strange et al. 2004),[5] it is likely that phonological considerations may also play a role in how L2 vowels are identified with L1 vowels, as predicted by the PAM-L2 (Section 2.3.3). In most cases phonetic and phonological comparisons will lead to the same result,

[5] Acoustic proximity in these studies has generally been measured in terms of distance in F_1 and F_2, but there are limits to estimating acoustic proximity in these terms, since F_1 and F_2, though sufficient as acoustic cues for distinguishing most vowels, are not the only determinants of vowel quality. Thus, it should be noted that inclusion of additional acoustic dimensions (especially f_0 and F_3, as well as temporal trajectories of these frequency components) would give a fuller picture of acoustic proximity between vowels and may help account for perceptual assimilations to an L1 vowel that is not the closest to an L2 vowel as measured on the basis of F_1 and F_2 alone.

but occasionally they will conflict. For instance, if it is assumed that perceptual linkage of L1 and L2 vowels is determined on the basis of acoustic proximity in F_1 and F_2 alone, English /o/ for female S&SW talkers is predicted to be linked to Korean /ɨ/, the acoustically closest Korean vowel in these terms, and thus over time to drift forwards and upwards in the vowel space towards Korean /ɨ/. On the other hand, if perceptual linkage is determined on the basis of phonological correspondence, English /o/ is predicted to be linked to Korean /o/, the corresponding mid back rounded vowel, and therefore to drift backwards in the vowel space toward Korean /o/.

Regardless of whether the cross-language vowel linkages are based on phonetics or phonology, though, the establishment of such linkages is expected to result in assimilatory phonetic drift of L1 vowels, and the dialectal differences highlighted in Section 5.1.1 lead to a clear prediction in this regard for an L1 vowel that is unambiguous in terms of L2 linkage—namely, English /ɑ/. On both phonological and phonetic grounds, English /ɑ/ should be linked to Korean /a/: Korean /a/ is the only low vowel in the Korean inventory and is, moreover, the closest Korean vowel in F_1 x F_2 phonetic space. This linkage predicts that, if L1 phonetic drift in vowels occurs at the level of individual vowels, phonetic drift in English /ɑ/ will show cross-dialectal variation due to its different starting position relative to Korean /a/. The relatively back English /ɑ/ of female MidA and S&SW talkers is expected to drift forwards and downwards in the vowel space towards Korean /a/; for female SoCal talkers, too, English /ɑ/ is expected to drift forwards, as well as upwards. On the other hand, the relatively front English /ɑ/ of female NMidW talkers is expected to drift slightly backwards or not drift at all, as it is already very close to Korean /a/ (Figure 5.1). Meanwhile, the relatively back English /ɑ/ of male MidA and S&SW talkers is expected to drift upwards towards Korean /a/, while the relatively front English /ɑ/ of male NMidW talkers is expected to drift both backwards and upwards towards Korean /a/ (Figure 5.2). Thus, the assumption of phonetic drift via cross-language linkages between individual vowels predicts a dissociation between the pattern of drift in English /ɑ/ manifested by NMidW talkers and that manifested by MidA and S&SW talkers because English /ɑ/ is significantly more front in the speech of NMidW talkers.

In short, like phonetic drift in the production of English stop consonants, phonetic drift in the production of English vowels is predicted to occur in a rapid and assimilatory fashion (in accordance with the hypotheses of Chapter 2), and this drift is predicted to occur at a structurally higher level than the segmental level. Rather than between individual vowels or natural classes of vowels, phonetic drift in English vowel production is predicted to occur on a global level, in approximation to the overall F_1 and F_2 levels of the Korean vowel space. It follows that drift in English vowels is expected to be realized similarly in talkers of various English dialects, since the overall F_1 and F_2 levels of the English vowel space differ from those of the Korean vowel space in the same direction across dialects due to systematic differences between the vowel inventories of the two languages. Contrary to the hypothesis of vowel-to-vowel drift, then, a dissociation between the drift pattern of NMidW talkers and the drift patterns of MidA and S&SW talkers is not expected.

5.2 Methods

Vowel productions were collected from learner participants in the weekly production experiments described in Chapter 3 (Section 3.4). Each week, participants read aloud the same set of Korean stimuli and the same set of English stimuli, and their responses were recorded digitally (Sections 3.5–3.6). Participants' recordings were acoustically analyzed in the manner described in Chapter 3 (Section 3.7.3). The critical items subjected to analysis comprised the 16 items with a glottal fricative onset and the two items with an aspirated bilabial stop onset (i.e., Korean /p^{h}a/, English [p^{h}ɑt]). The onset and offset of the vowel in these items were demarcated as shown in Figures 3.6–3.7. Automated measurements of F_1 and F_2 were then extracted over an interval of the middle 50 ms of the vowel.

Following this automatic formant extraction, the data were inspected for outliers by vowel and formant, and potential errors were flagged. Spectrograms of all tokens were then individually inspected to check that the formant tracking was accurate. When the formant tracking was irregular or inaccurate, the analysis parameters were adjusted until tracking was smooth, and new measurements were extracted. If the formant tracking could not be made satisfactory via adjustment of the analysis parameters, then measurements were taken manually on an average spectrum of the middle 50 ms of the vowel.[6] The data presented in Section 5.3 are based on a total of approximately 44 English vowel tokens per learner per week (11 English vowels x 4 tokens/vowel). Tokens that were anomalous in some way (e.g., pronounced on a yawn, cough, sigh, or burp) or where the wrong item was produced (e.g., *heed* for *head*) were discarded.[7]

Because of gender differences in vowel formants (see, e.g., Whiteside 1998a,b, as well as Simpson 2009 for a more general review of phonetic differences between male and female speech), male and female talkers are analyzed separately below. Moreover, error bars that are presented in group figures (including Figures 5.1–5.2) represent in every case the average of individual participants' standard errors for that particular vowel, rather than standard errors calculated over the entire distribution. This calculation prevents formant disparities due to physiological differences between participants, especially between male and female participants, from inflating the displayed error, allowing for a more accurate representation of the magnitude of the average participant's error for each vowel.

[6]The need to resort to manual measurement was rare, occurring in 0.3% of L2 learners' English tokens and 0.2% of native Korean speakers' Korean tokens.

[7]These tokens were very few in number, amounting to 1.8% of L2 learners' English tokens and 0.4% of native Korean speakers' Korean tokens.

5.3 Results

5.3.1 Change in English Vowels

The results of acoustic analyses indicate few dramatic changes in the English vowel space of the L2 learner group over the duration of the study. The English vowels of the learner participants for Weeks 1, 3, and 5 are plotted in Figures 5.3 and 5.4 for the female and male participants, respectively. Weeks 2 and 4 are included in all the statistical analyses, but are omitted from these and following figures for clarity of presentation.

For female learners (Figure 5.3), there is a subtle, but significant overall raising of the English vowels: F_1 is found to decrease over time. The results of a repeated-measures analysis of variance (ANOVA) with within-subjects factors Vowel and Time show significant main effects on F_1 of both Vowel [$F(10, 125) = 615.01, p < 0.001$] and Time [$F(4, 38) = 4.40, p < 0.01$], though no interaction between the two factors [$F(40, 600) = 0.97, n.s.$]. A post-hoc examination of differences between weeks using Tukey's HSD test indicates that while the increases in F_1 between Weeks 1 and 2 and Weeks 3 and 4 are not significant, the decreases between Weeks 2 and 3 and Weeks 4 and 5 are each highly significant [$p < 0.001$] and result in a net F_1 decrease of approximately 17 Hz from Week 1 to Week 5 (Figure 5.5a).

On the other hand, there is no significant overall movement of female learners' English vowels in F_2 (Figure 5.5b). A repeated-measures ANOVA with the same factors as above shows a main effect on F_2 of Vowel [$F(10, 125) = 345.32, p < 0.001$], but no effect of Time [$F(4, 38) = 0.16, n.s.$]. A significant interaction between Vowel and Time [$F(40, 600) = 1.82, p < 0.01$] suggests that not all vowels fail to drift in F_2, and in fact repeated-measures ANOVAs by vowel show that there is a marginally significant effect of Time on F_2 in /ɛ/ and /o/ (Table 5.5). However, the results of Tukey's HSD test examining differences between weeks by vowel (Table 5.6) reveal no significant differences in F_2 between weeks for any vowel.

For male learners (Figure 5.4), there is no significant overall shift in the vowel space of the English vowels in either F_1 or F_2 (Figures 5.5c–5.5d). Although there is a slight net increase in mean F_1 (7 Hz) and a more substantial net increase in mean F_2 (26 Hz), neither increase is statistically significant. Repeated-measures ANOVAs show a main effect of Vowel on both F_1 [$F(10, 7) = 732.19, p < 0.001$] and F_2 [$F(10, 7) = 57.16, p < 0.001$], but no effect of Time on either F_1 [$F(4, 2) = 0.57, n.s.$] or F_2 [$F(4, 2) = 5.29, n.s.$]. Moreover, there is no interaction between Vowel and Time for F_1 [$F(40, 80) = 0.61, n.s.$] or F_2 [$F(40, 80) = 0.72, n.s.$].

To examine the development of the English vowels in more detail, repeated-measures ANOVAs were also conducted for each vowel. The results of these analyses are summarized in Table 5.5 and are consistent with the initial observation of little change in most of the individual English vowels over the duration of the study. Time has a significant main effect on the acoustic realization of vowels in only two cases:

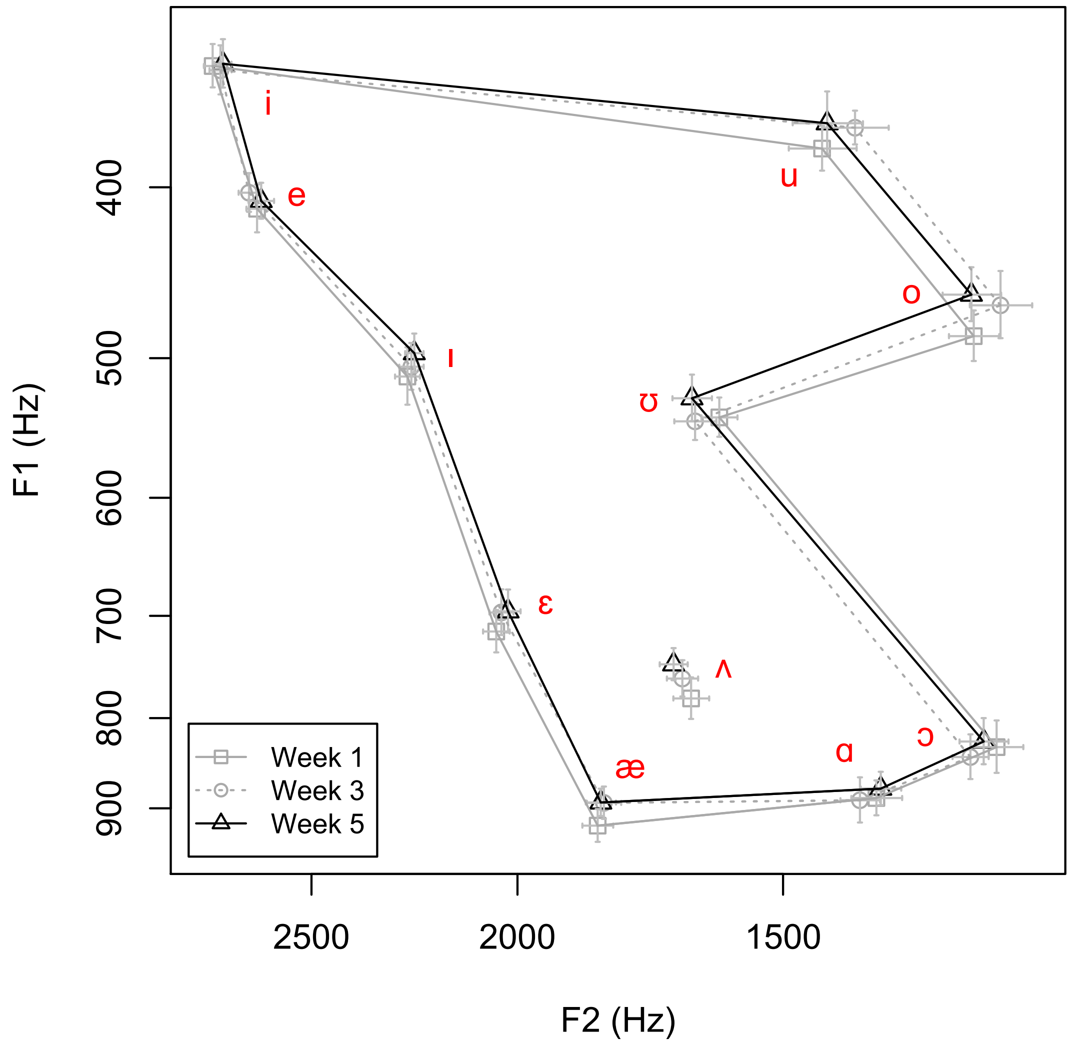


Figure 5.3: Mean F_1 by mean F_2 of English vowels over time as produced by female L2 learners of Korean. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.

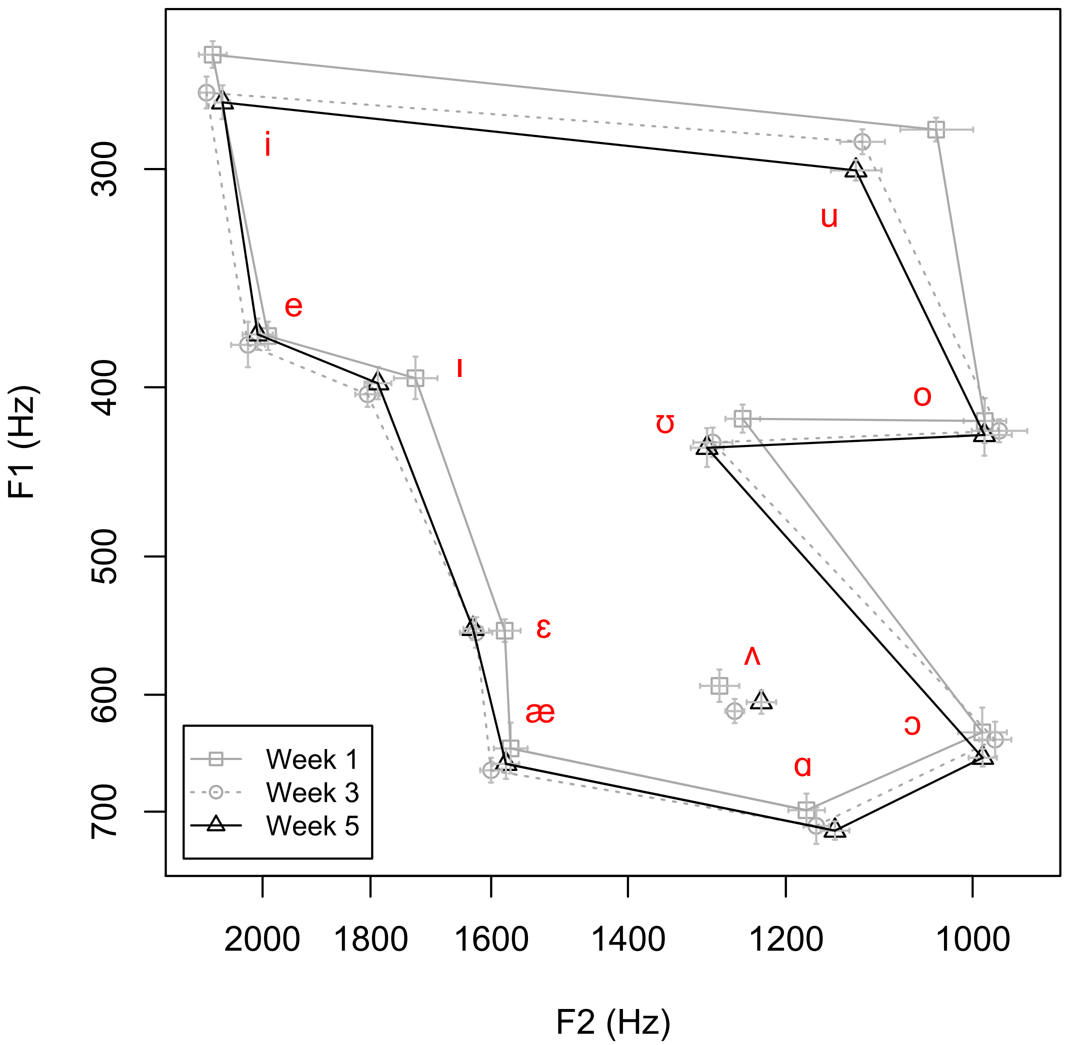
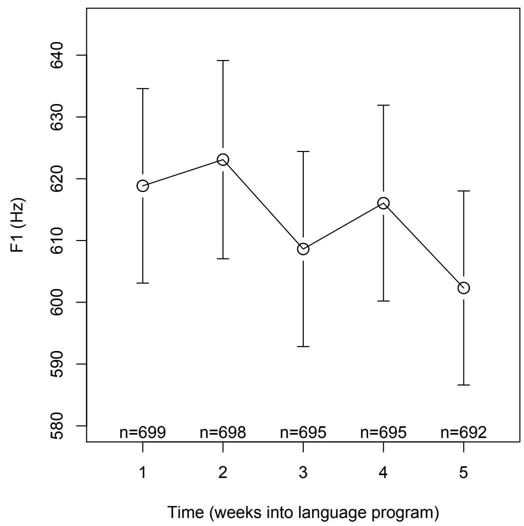
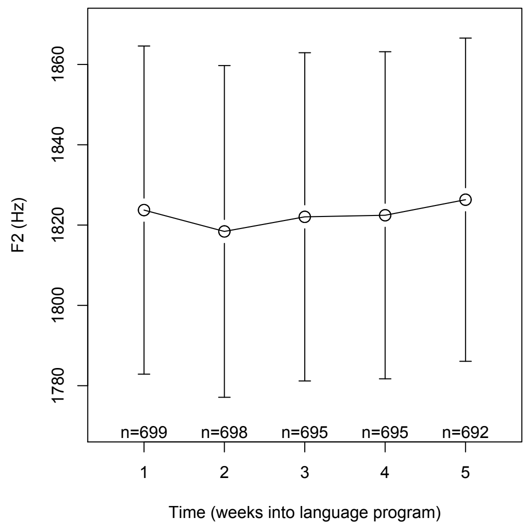


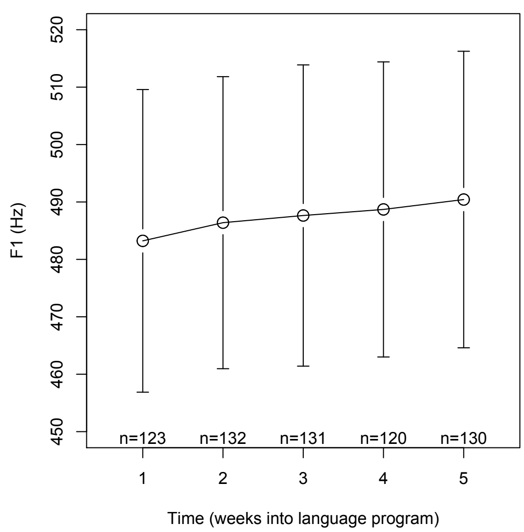
Figure 5.4: Mean F_1 by mean F_2 of English vowels over time as produced by male L2 learners of Korean. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.



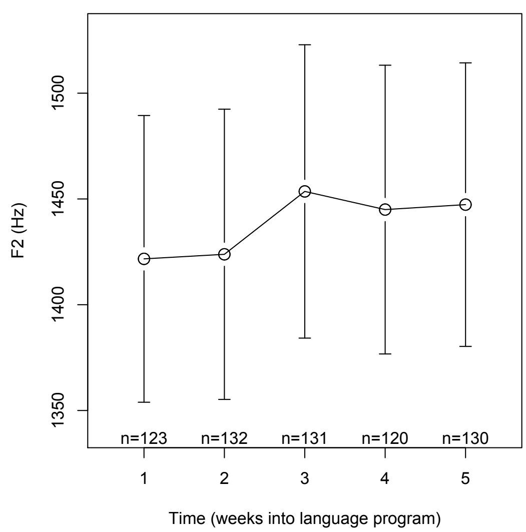
(a) Mean F_1, female learners



(b) Mean F_2, female learners



(c) Mean F_1, male learners



(d) Mean F_2, male learners

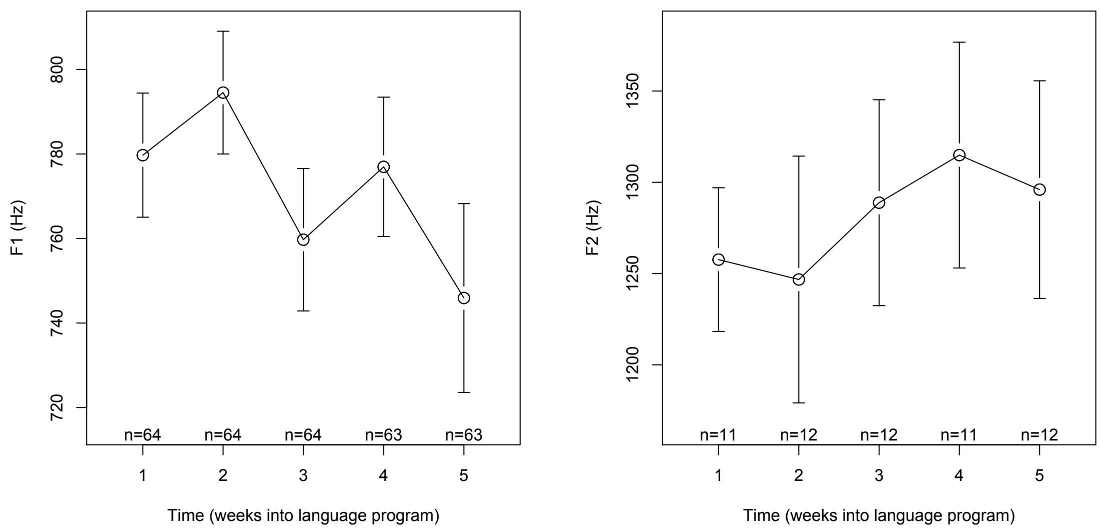
Figure 5.5: Mean F_1 and mean F_2 of the English vowel space over time, by formant and talker gender: (a) mean F_1, female learners; (b) mean F_2, female learners; (c) mean F_1, male learners; (d) mean F_2, male learners. Error bars indicate 95% confidence intervals.

Table 5.5: Results of repeated-measures analyses of variance testing for the effect of time on F_1 and F_2, by vowel and talker gender.

	F_1				F_2			
	female		male		female		male	
Vowel	$F_{(4,60)}$	p	$F_{(4,8)}$	p	$F_{(4,60)}$	p	$F_{(4,8)}$	p
/i/	0.15	*n.s.*	2.24	*n.s.*	1.40	*n.s.*	0.72	*n.s.*
/ɪ/	2.07	< 0.1	0.45	*n.s.*	1.01	*n.s.*	1.01	*n.s.*
/e/	1.10	*n.s.*	0.73	*n.s.*	0.90	*n.s.*	0.29	*n.s.*
/ɛ/	0.86	*n.s.*	0.05	*n.s.*	2.32	< 0.1	0.86	*n.s.*
/æ/	2.13	< 0.1	0.28	*n.s.*	0.95	*n.s.*	0.59	*n.s.*
/u/	0.86	*n.s.*	0.85	*n.s.*	1.03	*n.s.*	1.53	*n.s.*
/ʊ/	1.49	*n.s.*	1.87	*n.s.*	1.61	*n.s.*	5.30	< 0.05
/o/	1.22	*n.s.*	0.38	*n.s.*	2.36	< 0.1	0.32	*n.s.*
/ɑ/	0.61	*n.s.*	0.42	*n.s.*	1.89	*n.s.*	0.57	*n.s.*
/ɔ/	0.59	*n.s.*	0.66	*n.s.*	1.19	*n.s.*	0.73	*n.s.*
/ʌ/	4.95	< 0.01	1.48	*n.s.*	1.27	*n.s.*	0.66	*n.s.*

Table 5.6: Results of Tukey's HSD test examining differences in F_2 between weeks for female L2 learners, by vowel. Units are Hertz. Positive differences indicate an increase from the first week to the following week, negative differences a decrease.

	Weeks 1–2		Weeks 2–3		Weeks 3–4		Weeks 4–5	
Vowel	diff.	p	diff.	p	diff.	p	diff.	p
/i/	+5	*n.s.*	−29	*n.s.*	+8	*n.s.*	−15	*n.s.*
/ɪ/	+11	*n.s.*	−22	*n.s.*	−3	*n.s.*	−2	*n.s.*
/e/	−2	*n.s.*	+26	*n.s.*	−15	*n.s.*	−21	*n.s.*
/ɛ/	+8	*n.s.*	−23	*n.s.*	−18	*n.s.*	+7	*n.s.*
/æ/	+2	*n.s.*	−13	*n.s.*	+27	*n.s.*	−17	*n.s.*
/u/	−53	*n.s.*	−2	*n.s.*	−8	*n.s.*	+53	*n.s.*
/ʊ/	+19	*n.s.*	+25	*n.s.*	0	*n.s.*	+9	*n.s.*
/o/	−64	*n.s.*	+29	*n.s.*	−15	*n.s.*	+49	*n.s.*
/ɑ/	+20	*n.s.*	+2	*n.s.*	+11	*n.s.*	−39	*n.s.*
/ɔ/	+16	*n.s.*	+18	*n.s.*	−6	*n.s.*	−7	*n.s.*
/ʌ/	+18	*n.s.*	−3	*n.s.*	−4	*n.s.*	+19	*n.s.*



(a) Mean F_1, female learners' /ʌ/

(b) Mean F_2, male learners' /ʊ/

Figure 5.6: Cases of significant phonetic drift in individual English vowels: (a) mean F_1, female learners' /ʌ/; (b) mean F_2, male learners' /ʊ/. Error bars indicate 95% confidence intervals.

F_1 of females' /ʌ/ and F_2 of males' /ʊ/. Female learners' /ʌ/ decreases in F_1 by approximately 34 Hz from Week 1 to Week 5, while male learners' /ʊ/ increases in F_2 by approximately 38 Hz over the same time period (Figure 5.6). As mentioned above, there are also marginal effects of Time on F_2 of females' /ɛ/ and /o/, as well as on F_1 of females' /ɪ/ and /æ/. Note, however, that with Bonferroni correction for multiple tests, only the decrease in F_1 of females' /ʌ/ can be considered significant.

The main finding of the statistical analyses described above was a decrease in female learners' overall F_1, a result that was found to be statistically significant for the vowel /ʌ/. Male learners, in contrast, did not manifest any significant change in their English vowel space, and by-vowel analyses largely failed to show significant phonetic drift in individual English vowels. These findings should be interpreted with caution, however, as they comprise mostly null results. Perhaps so many null results obtained because participants manifested opposite patterns that canceled each other out, or because there was simply too much variance in formant frequencies (due to physiological differences among participants) at all time points for trends to reach the level of significance. In these cases, the null results would belie meaningful drift patterns for the whole vowel space or for individual vowels.

In order to factor out inter-speaker differences in a more fine-grained way, linear mixed-effects models (see, e.g., Pinheiro and Bates 2000; Baayen 2008:241–299;

Johnson 2008:237–247) were fit to the formant data, with Participant as a random effect and Vowel, Gender, Time, and their two- and three-way interactions as fixed effects. The predictors found to be significant in a model fit to the F_1 data with all effects entered are entirely consistent with the ANOVA results described above. Vowel $[F(10, 3988) = 5664.82, p < 0.001]$ and Time $[F(4, 3988) = 10.42, p < 0.001]$ each have significant effects on F_1. In addition, Gender has a strong effect on F_1: being male decreases F_1 relative to being female $[\beta = -84.42, t(17) = -3.10, p < 0.01]$, as expected. The interaction of Vowel and Gender $[F(10, 3988) = 87.71, p < 0.001]$ and the interaction of Gender and Time $[F(4, 3988) = 4.66, p < 0.01]$ improve the model further. On the other hand, the Vowel x Time interaction $[F(40, 3988) = 1.19, n.s.]$ and the Vowel x Gender x Time interaction $[F(40, 3988) = 0.35, n.s.]$ do not significantly improve the model.

Given these significant predictors and the focus on the effect of time, three kinds of F_1 models were built: those with only one fixed effect (Vowel, Gender, or Time), those with all fixed effects excluding Time (Vowel, Gender, Vowel x Gender), and those with all fixed effects including Time (Vowel, Gender, Time, Vowel x Gender, Gender x Time). In every case, the robustness of the model was measured via cross-validation (Johnson 2008:238–240): the same type of model was built 1000 times based upon a random subset of 85% of the data, and its predictions were then tested against the actual values for the dependent variable in the remaining 15% of the data.

Models fit to the F_1 data with one fixed effect indicate that of Vowel, Gender, and Time, Vowel has, unsurprisingly, the strongest effect on F_1. The 95% confidence interval for the percentage of variance accounted for (r^2) in models with only Vowel as a fixed effect is 0.914–0.934. In contrast, the 95% confidence interval for r^2 in models with only Gender or Time as a fixed effect is 0.034–0.090. In short, vowel identity is the most informative predictor of F_1 in a particular vowel production, accounting by itself for over 90% of variance in F_1.

Nevertheless, inclusion of Gender as well as Time significantly improves upon Vowel-only models of F_1. When Gender and the interaction between Vowel and Gender are added to Vowel as fixed effects, mean r^2 increases by approximately 0.013 from 0.924 to 0.937 $[t(1949) = 63.35, p < 0.001]$, while the root mean square (RMS) of the residual errors decreases from 57.65 to 52.47 $[t(1992) = 65.56, p < 0.001]$. Finally, when Time and the interaction between Gender and Time are added to the model, mean r^2 increases slightly more to 0.938 $[t(1993) = 2.52, p < 0.01]$ and the RMS of the residuals decreases further to 52.22 $[t(1993) = 3.19, p < 0.001]$. Thus, the effect of Time on F_1, while significant, is much smaller than the effects of Vowel and Gender.

Importantly, in neither the ANOVAs nor the mixed-effects modeling did a significant interaction obtain between Vowel and Time, suggesting that the decreasing trend in females' F_1 is not specific to the vowel /ʌ/, the only vowel for which the effect was found to be significant in ANOVAs. Indeed, when mean F_1 of female learners is examined by vowel, it is found that with the exception of /ɔ/, all of the vowels show a trend of decreasing F_1 over time (Figure 5.7). However, for nearly

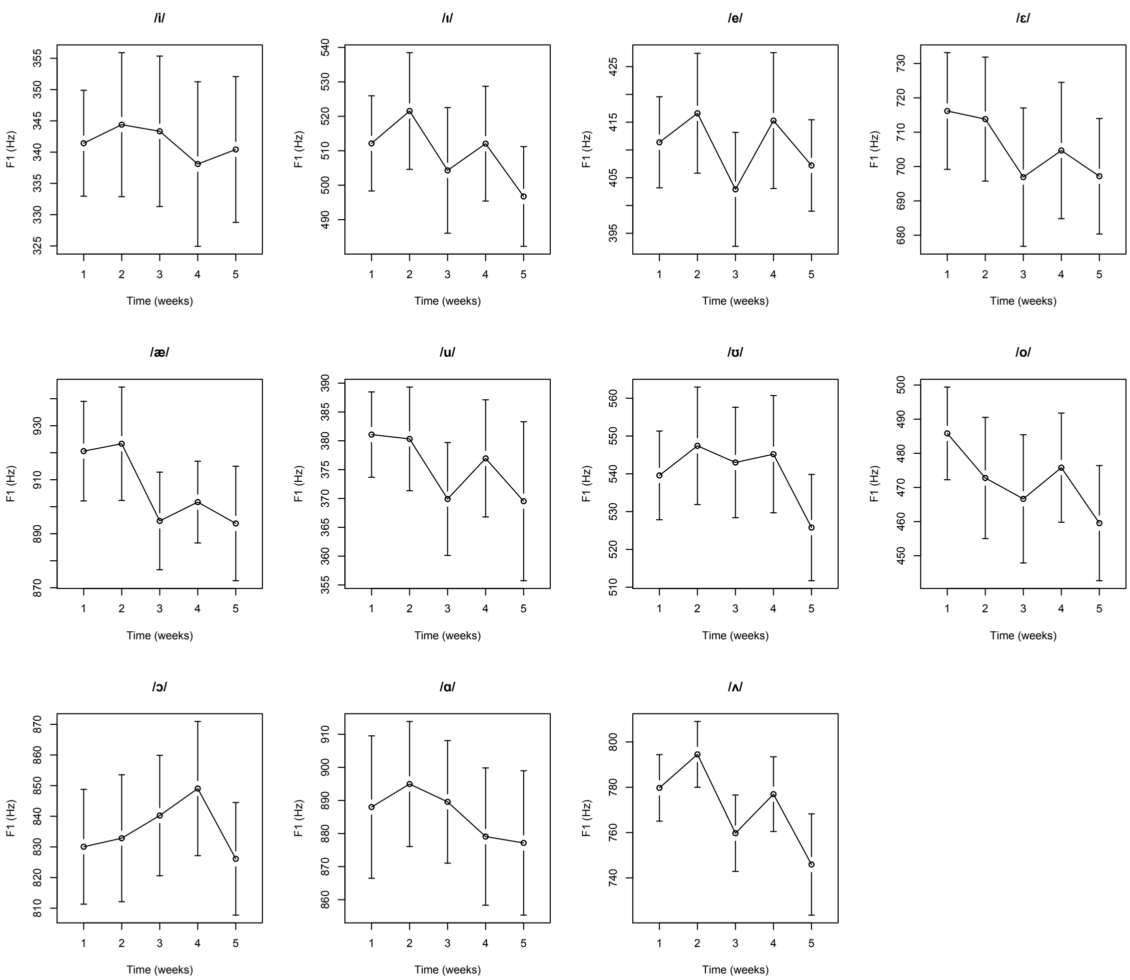


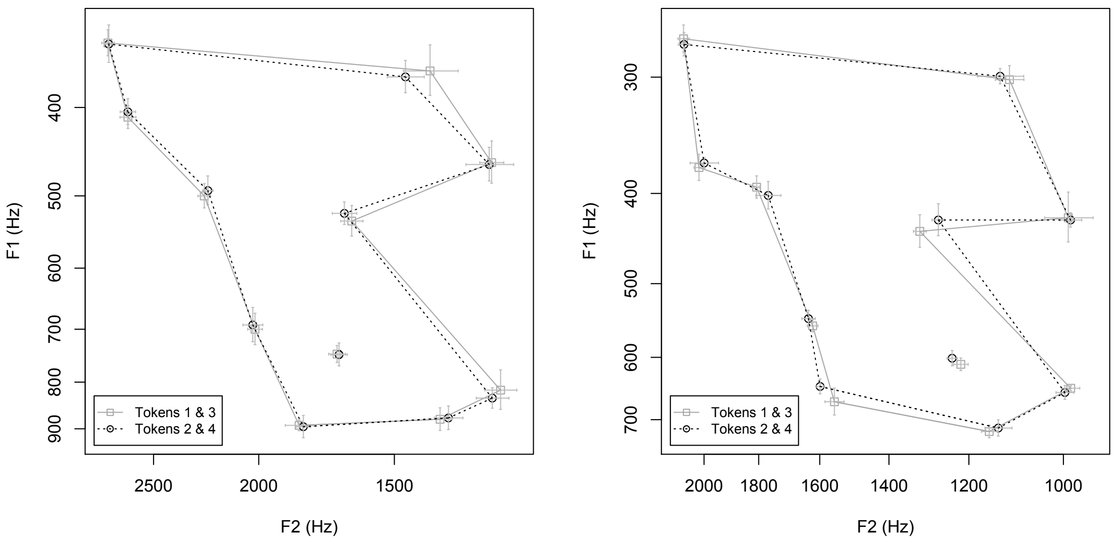
Figure 5.7: Mean F_1 of female L2 learners' English vowels over time, by vowel. Plots are labeled with the standard IPA transcriptions of the vowels. Error bars indicate 95% confidence intervals.

every vowel there is a large degree of variance at each time point, which appears to have prevented the trend from reaching significance in ANOVAs (Table 5.5). Thus, to factor out this inter-speaker variability, mixed-effects models of F_1 were also built by vowel, with Participant as a random effect and Gender, Time, and the Gender x Time interaction as fixed effects. These models indicate that Time is a significant predictor of F_1 not only in /ʌ/ [$F(4, 347) = 9.82, p < 0.001$], but also in /ɪ/ [$F(4, 349) = 3.87, p < 0.01$], /e/ [$F(4, 348) = 2.56, p < 0.05$], /æ/ [$F(4, 345) = 3.97, p < 0.01$], /ʊ/ [$F(4, 344) = 3.15, p < 0.05$], and /o/ [$F(4, 345) = 2.46, p < 0.05$]. Time is a marginally significant predictor of F_1 in /ɛ/ as well [$F(4, 348) = 2.04, p < 0.1$]. These findings thus suggest that the significant decrease in females' overall F_1 is a robust result holding of the vowel space in general, rather than just of the vowel /ʌ/.

A linear mixed-effects model was also fit to the F_2 data, with the same random and fixed effects as above. The predictors found to be significant in this model are also consistent with the ANOVA results. Vowel has a significant effect on F_2 [$F(10, 3988) = 5941.00, p < 0.001$], as does Gender: being male again decreases formant frequencies relative to being female [$\beta = -683.45, t(17) = -11.08, p < 0.001$]. However, in contrast to the F_1 model, Time does not have a significant effect on F_2 [$F(4, 3988) = 0.19, n.s.$], nor do any of its interactions: Vowel x Time [$F(40, 3988) = 1.12, n.s.$], Gender x Time [$F(4, 3988) = 0.96, n.s.$], or Vowel x Gender x Time [$F(40, 3988) = 0.57, n.s.$]. On the other hand, the interaction of Vowel and Gender does have a significant effect on F_2 [$F(10, 3988) = 79.38, p < 0.001$].

Similar to models of F_1, models of F_2 with one fixed effect indicate that Vowel has the strongest effect on F_2. However, inclusion of Gender significantly improves upon Vowel-only models. When Gender and the Vowel x Gender interaction are included in the model, mean r^2 increases by approximately 0.012 from 0.929 to 0.941 [$t(1949) = 63.35, p < 0.001$], while the RMS of the residuals decreases from 144.27 to 132.41 [$t(1949) = 63.35, p < 0.001$]. In contrast, the inclusion of Time in the model fails to improve its performance. In fact, it slightly worsens the predictions of the model, resulting in a non-significant decrease in mean r^2 of 0.0003 [$t(1998) = -1.33, n.s.$] and a non-significant increase in the RMS of the residuals of 0.10 [$t(1998) = -0.49, n.s.$].

To summarize, after accounting for the effects of vowel, gender, and participant in linear mixed-effects models of vowel formants, time is still found to have an effect on F_1 (but not F_2) of the English vowels, though males and females differ in terms of the effect that both vowel and time have on their produced F_1. The gender difference with respect to the effect of vowel can be seen in comparing Figure 5.3 with Figure 5.4: although the male and female vowel spaces are organized similarly, the vowels differ between the genders in terms of their precise location in the F_1 dimension relative to a given reference point. For example, /e/ lies much farther below /i/ for male learners than for female learners, such that it is almost even with /ɪ/. In addition, /æ/ is basically even with /ɑ/ for female learners, while it is higher than /ɑ/ for male learners; conversely, /o/ is even with /ʊ/ for male learners, while it is higher than /ʊ/ for female learners. The gender difference with respect to the effect of time was seen



(a) Female learners

(b) Male learners

Figure 5.8: Variability of English vowel production in Week 5, by talker gender: (a) female L2 learners; (b) male L2 learners. In each plot, mean F_1 is plotted by mean F_2 on a logarithmic scale. Means over the first half of productions (Tokens 1 and 3) are represented with squares connected by solid gray lines; means over the second half of productions (Tokens 2 and 4), with circles connected by dotted black lines. Error bars indicate ±1 standard error about the mean.

in Figure 5.5: female learners showed a significant decline in F_1 over time, whereas male learners showed a small, non-significant increase.

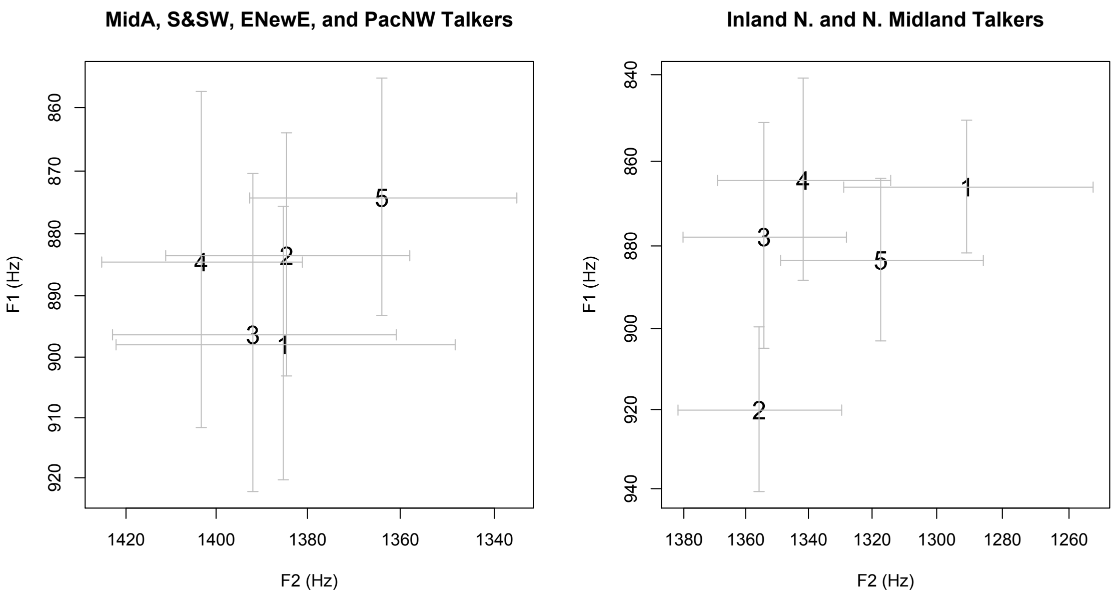
Thus, although time was found to have a significant effect on English vowel production, the effect was quite subtle. In order to check whether the observed shifts were in fact different from normal speech variability one would expect to find at a given point in time, the consistency of learners' English vowel production was examined at the last measurement point—Week 5 of the language program. This was done by splitting the data for Week 5 into two halves of non-consecutive tokens (i.e., Tokens 1 and 3, and Tokens 2 and 4). Separate means were calculated for each half of the data, and these are plotted in Figure 5.8. For both female and male learners, there is some variability in the production of one or more vowels at this time point—in particular, /u/ for females (Figure 5.8a), and /æ/ and /ʊ/ for males (Figure 5.8b). However, in general this analysis indicates that, at one point in time, learners were very consistent in their English vowel production, suggesting that the directional drift observed in their English vowels over time was not simply an artifact of normal speech variability.

5.3.2 Dialectal and Individual Variation

In Section 5.1.2, it was predicted that phonetic drift in English vowels would occur on a global level, resulting in a systemic shift of the English vowel space approximating the different spectral norms of the Korean vowel space. Due to the global nature of this drift, it was predicted that, despite some clear dialectal differences in the organization of the English vowel space, phonetic drift in English vowels would be realized similarly for talkers of different dialects, since phonetic discrepancies at a global level between the English vowel space of various dialects and the Standard Korean vowel space are similar, arising from systematic differences between the two languages' vowel inventories. In this section, we first consider differences in drift patterns between speakers of different dialects, focusing on the more numerous female learners specifically. In particular, the possibility of dialectal differences in phonetic drift of English /ɑ/, due to differences in the directionality of potential influence from Korean /a/, is thoroughly investigated. We then turn to inter-speaker variation in shifts of the English vowel space as a whole.

Based upon their hometowns, residential history, and the geographic origins of their parents, the majority of the learner participants in the present study can be expected to have a vowel space that contains key features of the vowel spaces described in Section 5.1.1. Three learners (LF31, LF37, LF55) may be identified as Mid-Atlantic talkers and three (LF03, LF16, LF32) as Southern or Southwestern talkers. The remaining talkers are from the Inland North (LM23, LM44, LF46), North Midland (LF18, LF19, LF28, LF54), Eastern New England (LF05, LF25, LF47), or Pacific Northwest (LF06, LM13, LF29). No talkers have geographic origins squarely in the northern Midwest or southern California. Nonetheless, the features of northern Midwest speech relevant to the realization of English /ɑ/—in particular, the raising and tensing of /æ/ and fronting of /ɑ/ resulting from the Northern Cities Shift—can be expected to be found in the dialects of the Inland North and North Midland, which are also affected by the Northern Cities Shift (Labov et al. 2006). In addition, features of southern California speech relevant to the production of English /ɑ/—especially the merger of /ɑ/ and /ɔ/—are also found in the dialect of the Pacific Northwest.

These facts, combined with the tendency of Eastern New England talkers to preserve a back /ɑ/ in words like *hot*, are suggestive of a dialectal divide between learners with a relatively back English /ɑ/ (i.e., Mid-Atlantic, South/Southwest, Eastern New England, and Pacific Northwest talkers) and learners with a relatively front English /ɑ/ (i.e., Inland North and North Midland talkers). Thus, if phonetic drift of English vowels occurred on the basis of vowel-to-vowel linkages to Korean vowels, it is expected that the relatively back English /ɑ/ of the former group would drift forwards in the vowel space towards the perceptually linked Korean /a/, while the relatively front English /ɑ/ of the latter group would drift slightly backwards or not at all, as it would have started only slightly forward of Korean /ɑ/. However, this sort of dissociation in drift patterns is not found. When learners' longitudinal production of English



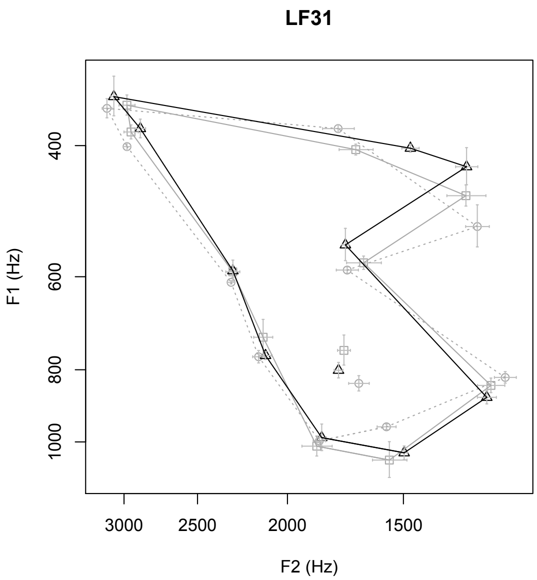
(a) Non-fronting dialects

(b) Fronting dialects

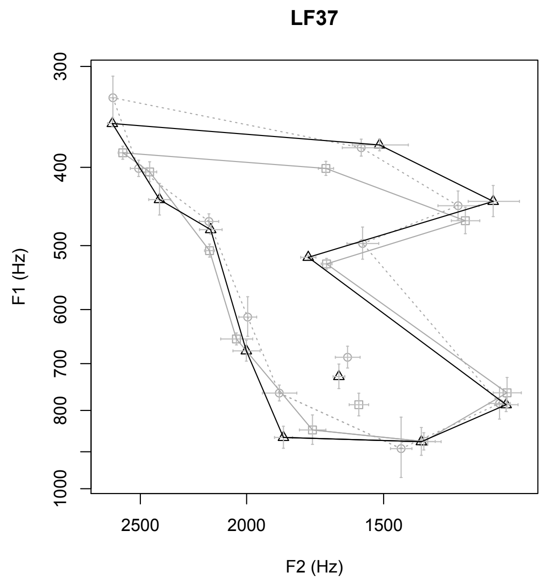
Figure 5.9: Mean F_1 by mean F_2 of English /ɑ/ over time in female learners, by dialect group: (a) the non-fronting dialects of the Mid-Atlantic, South/Southwest, Eastern New England, and Pacific Northwest; (b) the fronting dialects of the Inland North and North Midland. The scale of both axes is logarithmic. Means are plotted with numerical symbols corresponding to the week in the language program. Error bars indicate ±1 standard error about the mean.

/ɑ/ is examined by dialect group, the English /ɑ/ of neither group is found to drift in a clear direction over time (Figure 5.9). Rather than simply drifting forward, the English /ɑ/ of the non-fronting group drifts forward and then backward, ultimately landing behind its position in Week 1 (Figure 5.9a). Meanwhile, the English /ɑ/ of the fronting group also drifts forward and then backward, ultimately landing in front of its position in Week 1 (Figure 5.9b). These findings, therefore, fail to support the prediction of dialect-based disparities in the drift of English vowels arising from different configurations of the English vowels in relation to Korean vowels.

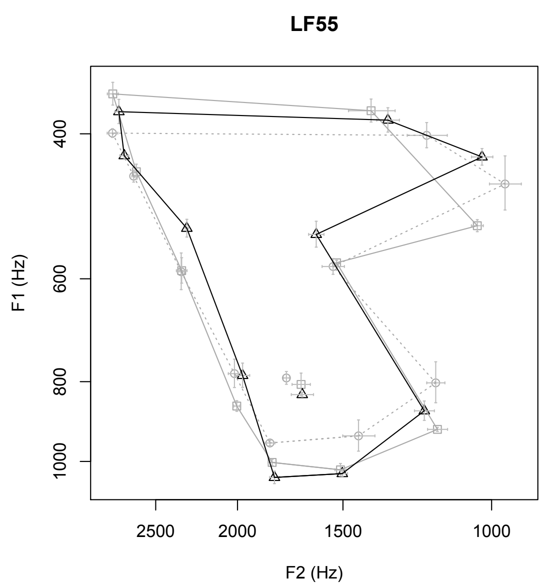
Inspection of the English vowel spaces of individual participants shows considerable variation among learners with respect to both the organization of the vowel space and patterns of drift in individual vowels. By-participant plots of mean F_1 by mean F_2 are provided below for the Mid-Atlantic group (Figure 5.10), South/Southwest group (Figure 5.11), Inland North group (Figure 5.12), North Midland group (Figure 5.13), Eastern New England group (Figure 5.14), and Pacific Northwest group (Figure 5.15). These individual plots are parallel to the group plots in Figures 5.3–5.4.



(a) Participant LF31

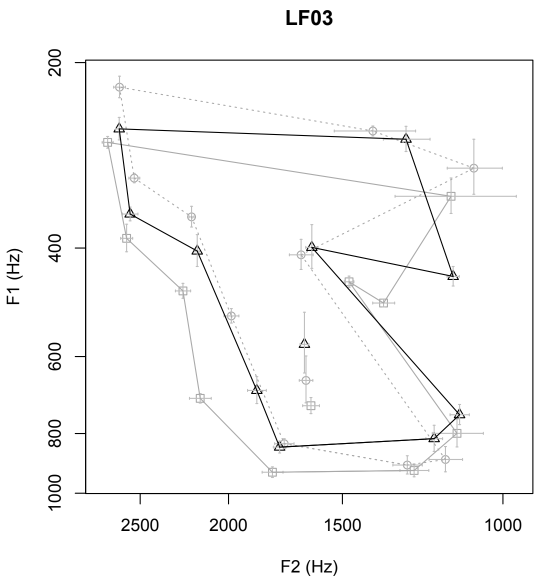


(b) Participant LF37

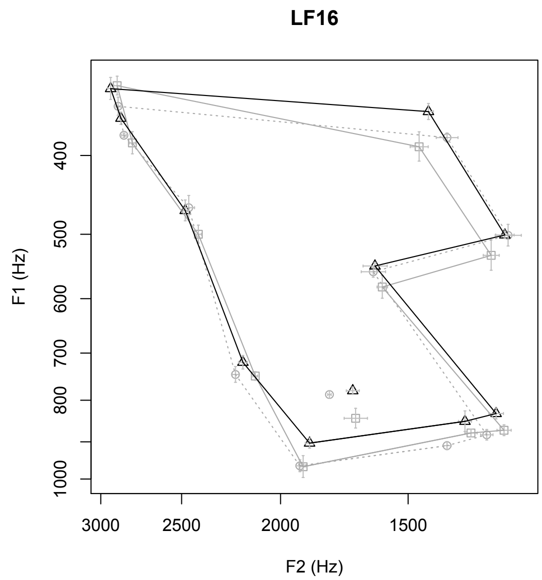


(c) Participant LF55

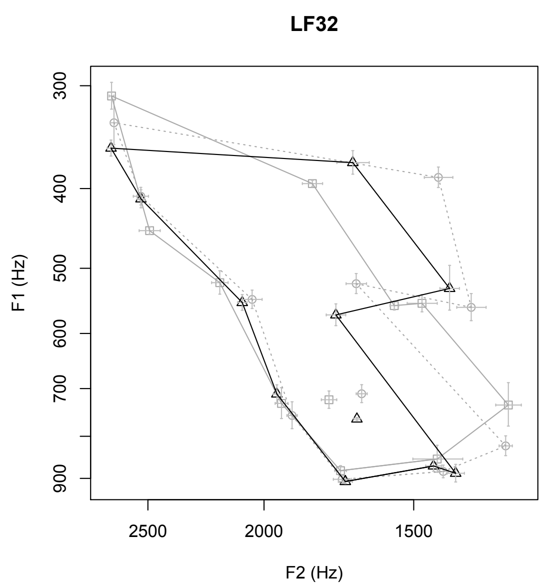
Figure 5.10: Mean F_1 by mean F_2 of English vowels over time as produced by talkers from the Mid-Atlantic. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.



(a) Participant LF03

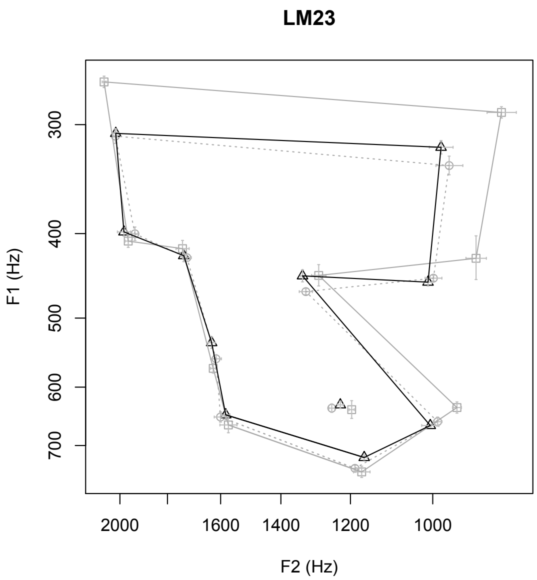


(b) Participant LF16

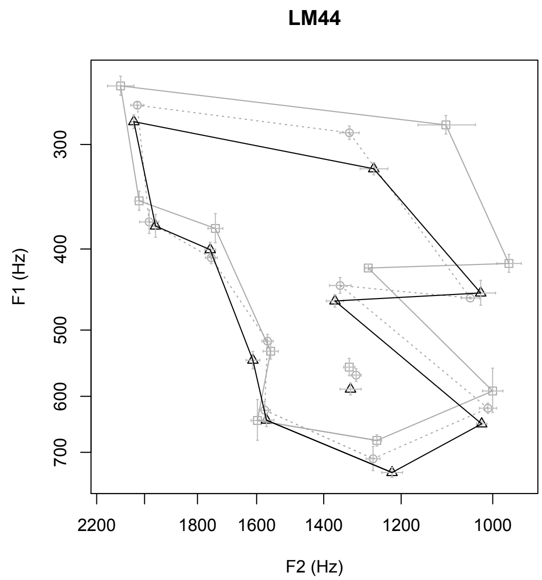


(c) Participant LF32

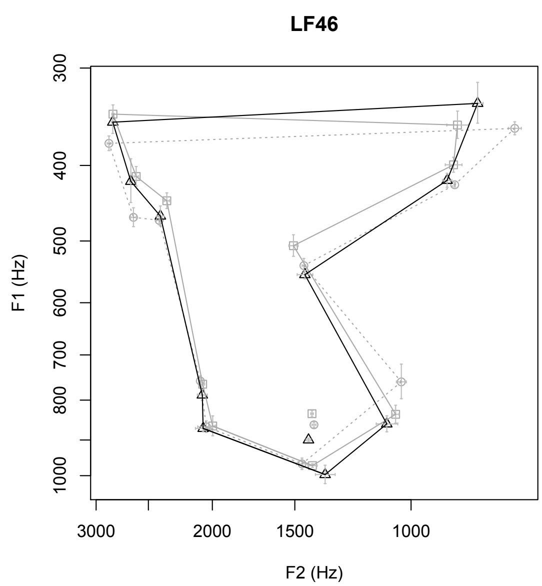
Figure 5.11: Mean F_1 by mean F_2 of English vowels over time as produced by talkers from the South/Southwest. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.



(a) Participant LM23

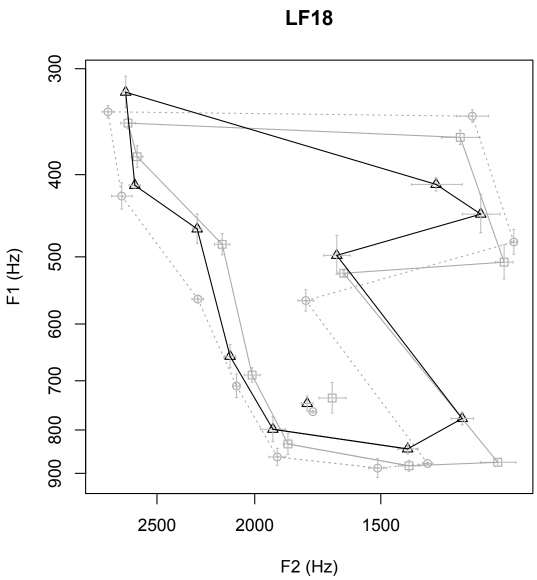


(b) Participant LM44

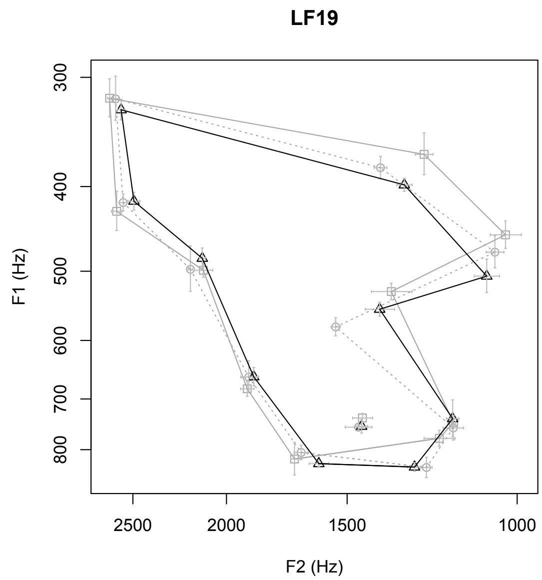


(c) Participant LF46

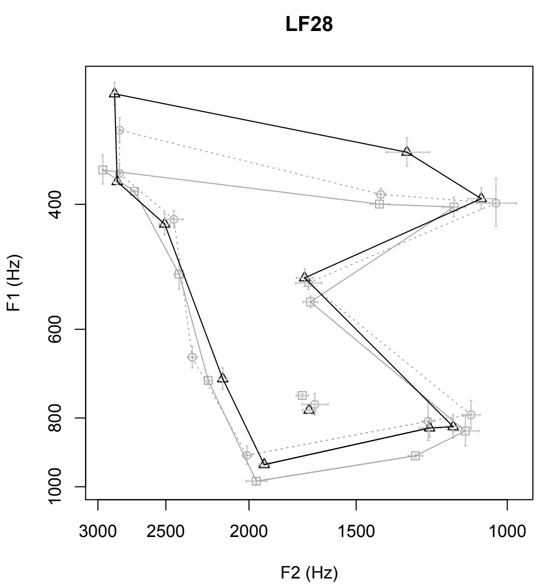
Figure 5.12: Mean F_1 by mean F_2 of English vowels over time as produced by talkers from the Inland North. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.



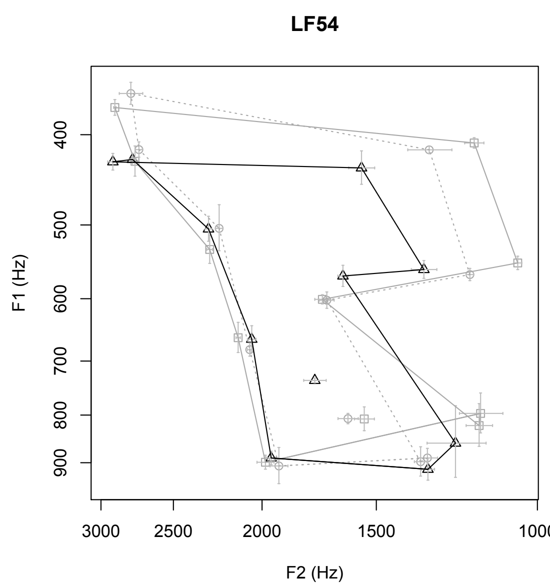
(a) Participant LF18



(b) Participant LF19

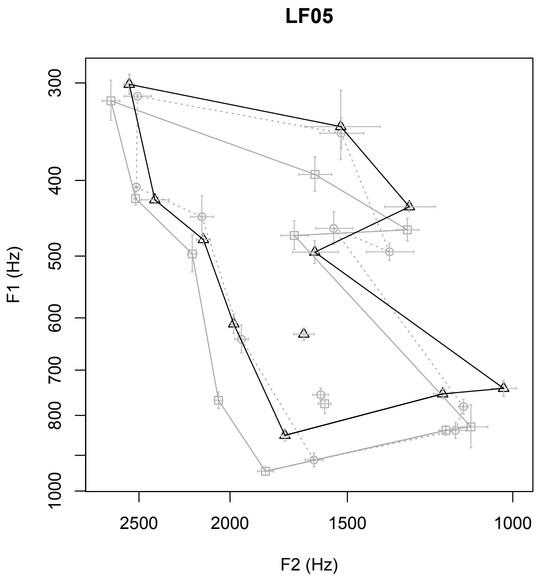


(c) Participant LF28

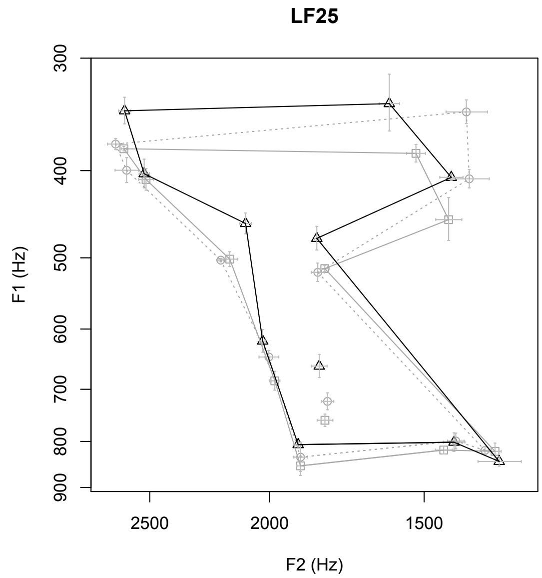


(d) Participant LF54

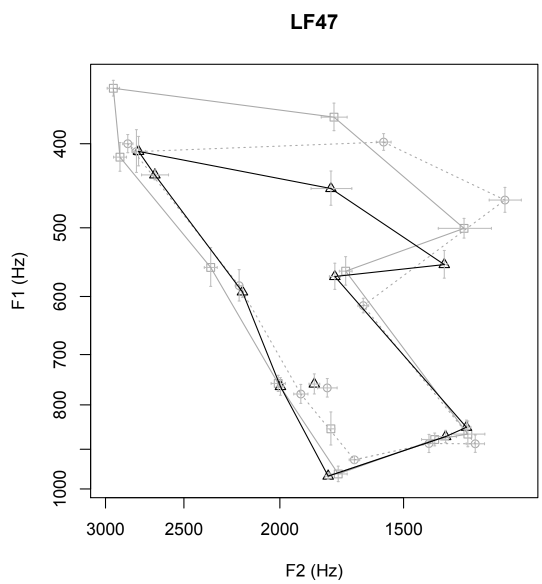
Figure 5.13: Mean F_1 by mean F_2 of English vowels over time as produced by talkers from the North Midland. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.



(a) Participant LF05

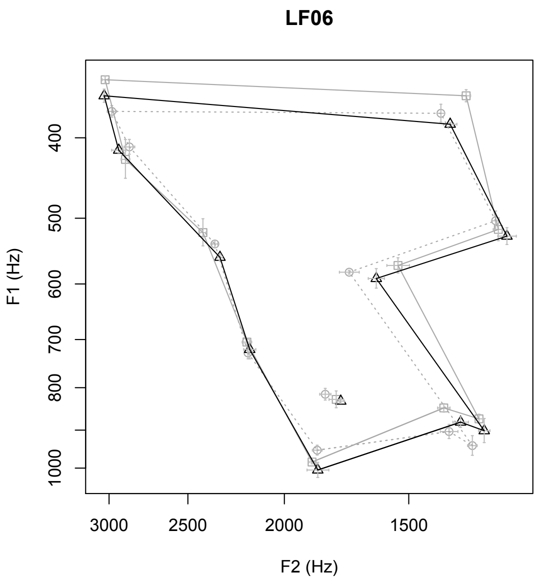


(b) Participant LF25

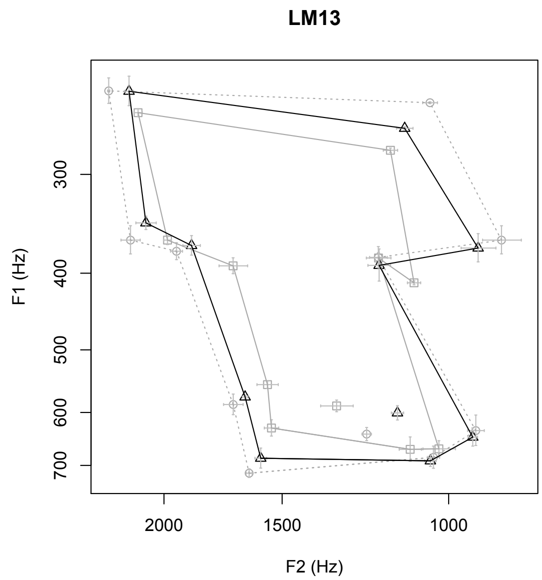


(c) Participant LF47

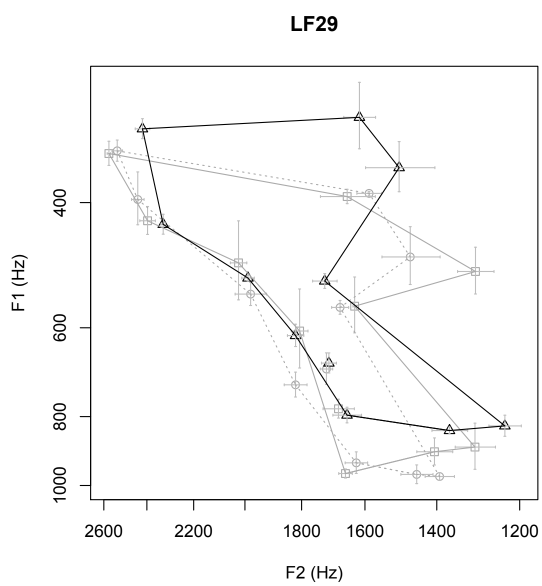
Figure 5.14: Mean F_1 by mean F_2 of English vowels over time as produced by talkers from Eastern New England. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ± 1 standard error about the mean.



(a) Participant LF06



(b) Participant LM13



(c) Participant LF29

Figure 5.15: Mean F_1 by mean F_2 of English vowels over time as produced by talkers from the Pacific Northwest. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 standard error about the mean.

The first way in which participants differ is in the shape of their English vowel space. Some learners (e.g., LF16, Figure 5.11b; LF25, Figure 5.14b) have a trapezoidal vowel space in which /u/ and /ɑ/ are both located relatively far back. Other learners (e.g., LF46, Figure 5.12c; LM23, Figure 5.12a) have a triangular vowel space in which /u/ is located relatively far back, but /ɑ/ is located relatively far front. Finally, other learners (e.g., LF32, Figure 5.11c; LF47, Figure 5.14c) have a parallelogram-shaped vowel space, in which /u/ is located relatively far front, but /ɑ/ is located relatively far back. These different configurations of the vowel space recall those documented in the studies of Peterson and Barney (1952), Hillenbrand et al. (1995), Yang (1996a), and Hagiwara (1997), which are compared in Figures 5.1–5.2. However, it should be noted that participants' regional affiliations do not fully determine the way they produce vowels in the experimental task. In particular, more participants than expected tend towards a classic trapezoidal vowel space, even though they might be predicted—based upon where they grew up—to show a vowel space shaped more like a triangle or parallelogram. In other words, dialectal differences in the organization of the vowel space appear to be attenuated in the citation register of speech elicited in the task.

The second way in which participants differ is in their drift patterns for individual vowels. There is considerable variation among participants with respect to the consistency of their productions of a given vowel at one time point: some participants (e.g., LF06, Figure 5.15a; LF46, Figure 5.12c) are quite consistent in their productions, while other participants (e.g., LF03, Figure 5.11a; LF05, Figure 5.14a) are more variable. Among both high- and low-variability participants the magnitude and direction of drift are found to differ. This is the case for a number of vowels. The vowel /æ/, for instance, hardly changes over time in the case of participant LF06 (Figure 5.15a), whereas it decreases in F_1 in the case of participant LF29 (Figure 5.15c). In addition, the vowel /u/ is found to increase primarily in F_2 for participant LF54 (Figure 5.13d), but to increase in both F_1 and F_2 for participant LM23 (Figure 5.12a); in contrast, for participants LF16 (Figure 5.11b) and LF28 (Figure 5.13c) /u/ decreases in F_1 and F_2. These inter-speaker differences notwithstanding, participants are generally not found to manifest directional drift patterns for individual vowels, instead showing either consistency of vowel productions over time or apparently random variation of vowel productions suggestive of a phonetically expansive vowel category. The one exception is the increase in F_1 of the vowel /ʌ/, a pattern that is manifested by half of the female participants for whom the effect was found (LF03, LF05, LF16, LF25, LF29, LF37, LF47, LF54). Thus, despite variation among participants, individual patterns are on the whole consistent with the group picture seen in Section 5.3.1: with the exception of /ʌ/, significant phonetic drift is not found for individual vowels.

The group trends in overall mean F_1 and F_2 levels (Figure 5.5) also hold true of the majority of individual learners. Consistent with the decrease in females' F_1 seen in Figure 5.5a, eight female learners (LF03, LF05, LF16, LF18, LF25, LF28, LF29,

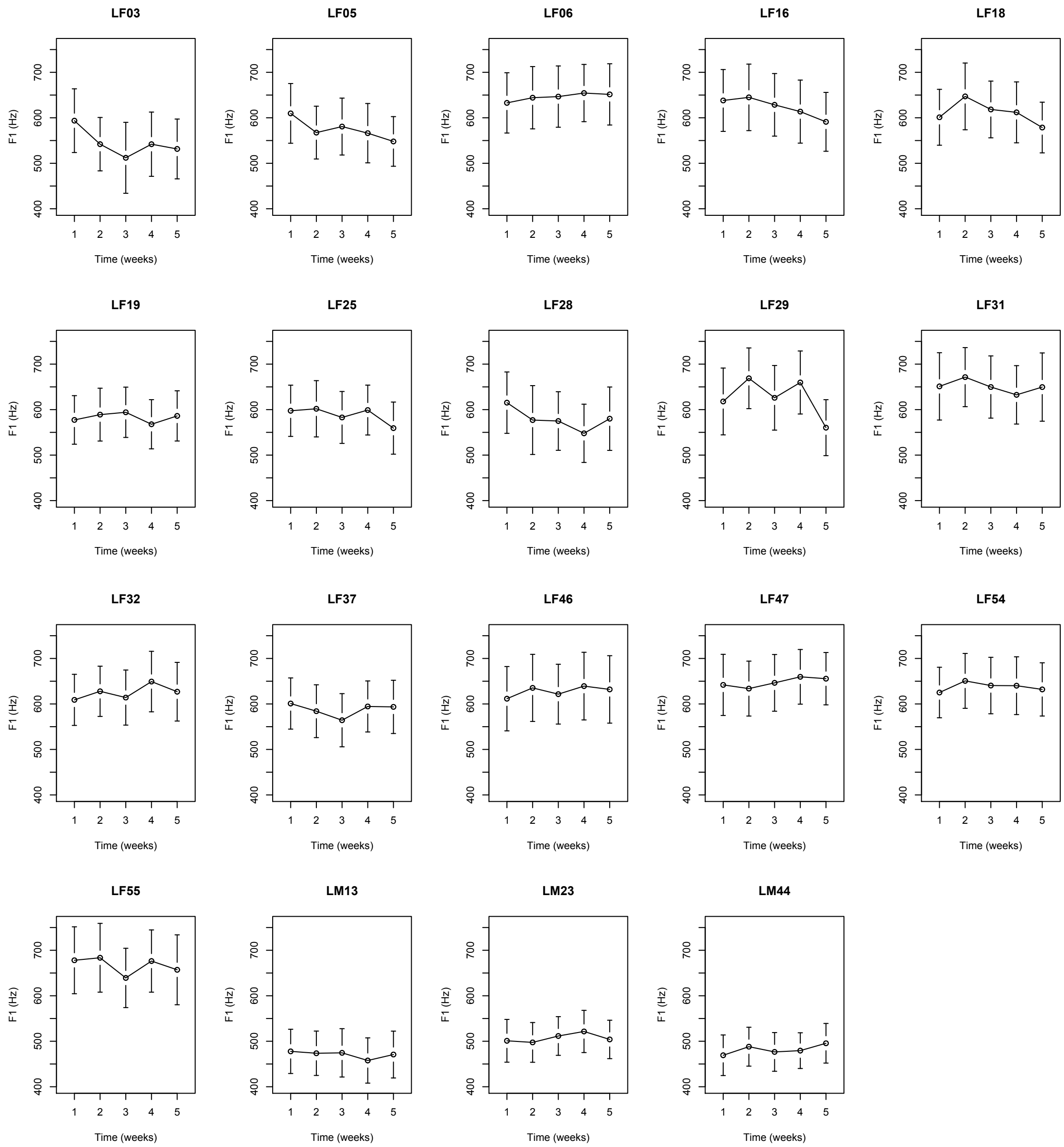


Figure 5.16: Mean F_1 of the English vowel space over time, by participant. Male learners LM13–LM44 are the last three shown. Error bars indicate 95% confidence intervals.

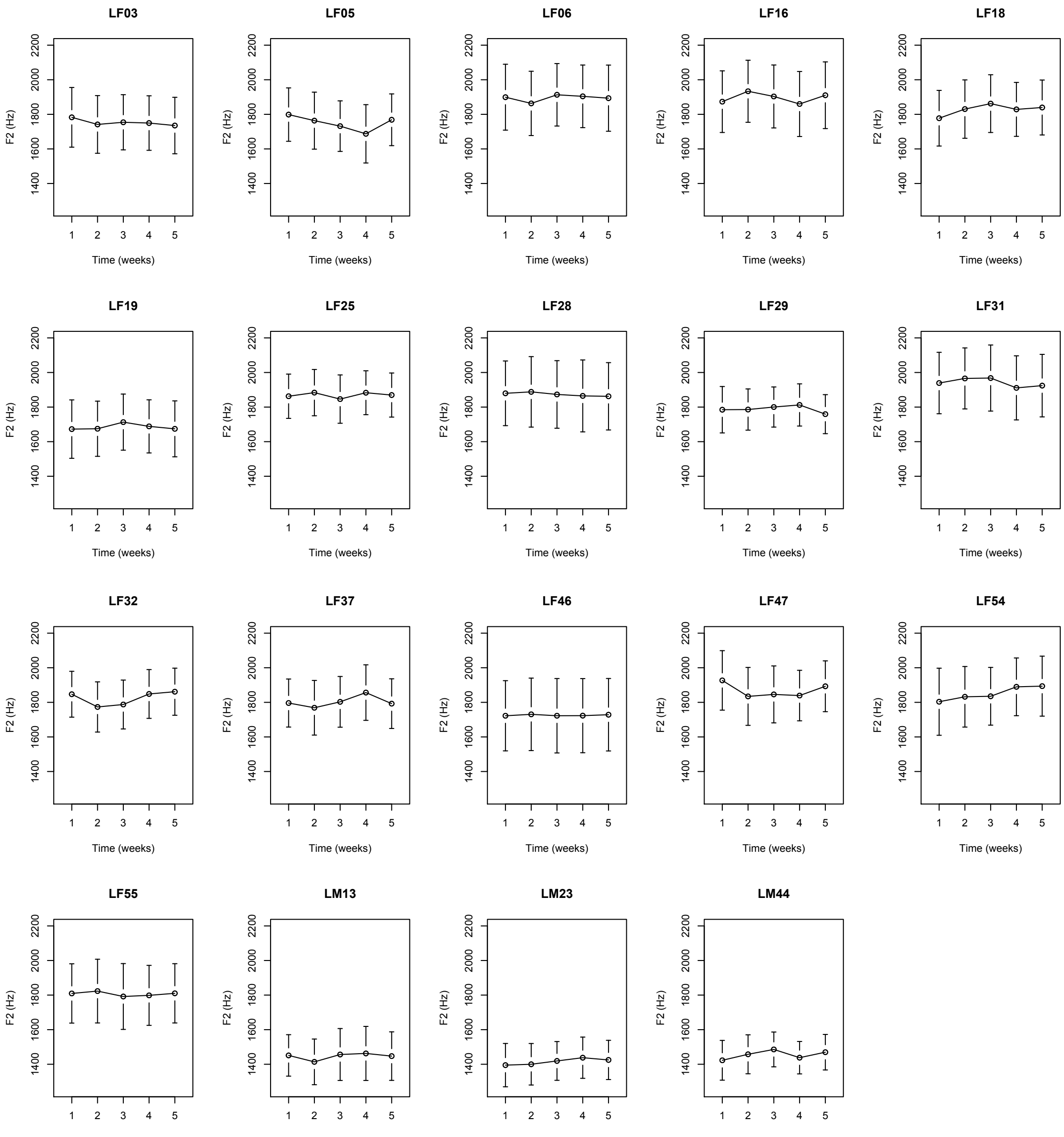


Figure 5.17: Mean F_2 of the English vowel space over time, by participant. Male learners LM13–LM44 are the last three shown. Error bars indicate 95% confidence intervals.

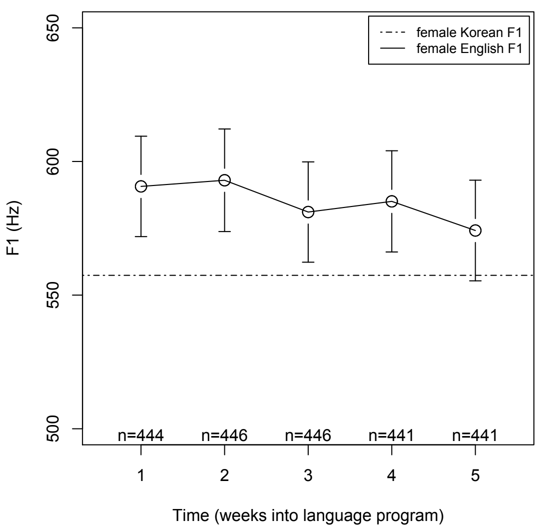
LF55) show a decrease in mean F_1, in comparison to four learners (LF19, LF37, LF31, LF54) whose F_1 fluctuates around a central tendency and four learners (LF06, LF32, LF46, LF47) who show the opposite trend (Figure 5.16). Of the three male learners, two (LM23, LM44) show a slightly increasing trend in mean F_1, whereas the other (LM13) shows a slight decrease. With respect to F_2, nearly all participants fail to show substantial change over time. While two participants (LF03, LF05) show a slight decrease in F_2 and five participants (LF18, LF32, LF54, LM23, LM44) show a slight increase, twelve participants (LF06, LF16, LF19, LF25, LF28, LF29, LF31, LF37, LF46, LF47, LF55, LM13) fluctuate around a central tendency (Figure 5.17).

5.3.3 Change in English Compared to Korean

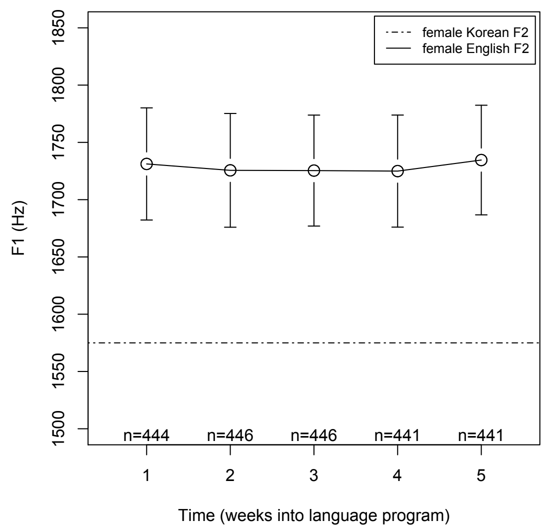
Comparisons of learners' English vowel production to Korean vowel norms suggest that the trends shown in Figure 5.5 are assimilatory in nature. In Figure 5.18, the mean F_1 and F_2 of the seven Korean vowels /i, ɛ, u, ɨ, o, ʌ, a/ (as produced by native Korean speaker participants in the present study) are plotted in comparison to the mean F_1 and F_2 of a parallel seven-vowel subset of the English vowels (/i, ɛ, u, ʊ, o, ʌ, ɑ/) produced by L2 learner participants.

With regard to the female L2 learners, the mean F_1 of female learners' English vowels drifts downward towards the lower mean F_1 of female Korean speakers' Korean vowels, such that the English vowels go from being significantly different from the Korean vowels in Week 1 to not being significantly different in Week 5 (Figure 5.18a). The F_1 difference between the two vowel spaces in Week 1 is small (approximately 33 Hz), but significant [$t(358) = 1.87, p < 0.05$], while the 17-Hz difference in Week 5 is not significant [$t(358) = 0.94, n.s.$]. On the other hand, no change is found in the case of F_2. The mean F_2 of female learners' English vowel space stays steady over time, at approximately 160 Hz above the mean F_2 of the female Korean vowel space from Week 1 to Week 5 (Figure 5.18b).

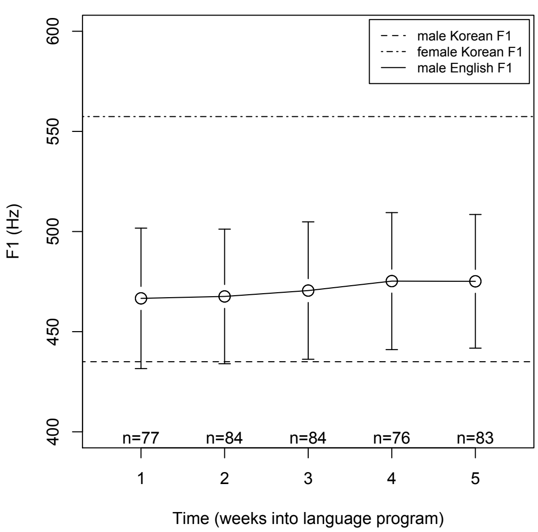
With regard to the male L2 learners, the mean F_1 of male learners' English vowels increases slightly, from 32 Hz above the mean F_1 of male Korean speakers' Korean vowels in Week 1 [$t(130) = 1.33, p < 0.1$] to 40 Hz above the male Korean vowels in Week 5 [$t(134) = 1.73, p < 0.05$]. At first glance, this might seem like a case of divergence between the L1 vowels and the L2 vowels, but in reality it is unlikely that the male Korean vowels served as the L2 model for the male L2 learners, since all of their Korean instruction during class (where they received the vast majority of their L2 exposure) came from female Korean speakers. In light of this fact, the most logical comparison to make is between the male English vowels and the female Korean vowels, which suggests that the slight increase in mean F_1 of the male English vowels is actually assimilatory to the higher mean F_1 of the female Korean vowels (Figure 5.18c). Likewise, while the mean F_2 of male English vowels diverges from the mean F_2 of male Korean vowels, increasing from 19 Hz above the Korean vowels in Week 1 to 45 Hz above the Korean vowels in Week 5, this F_2 increase brings the mean F_2



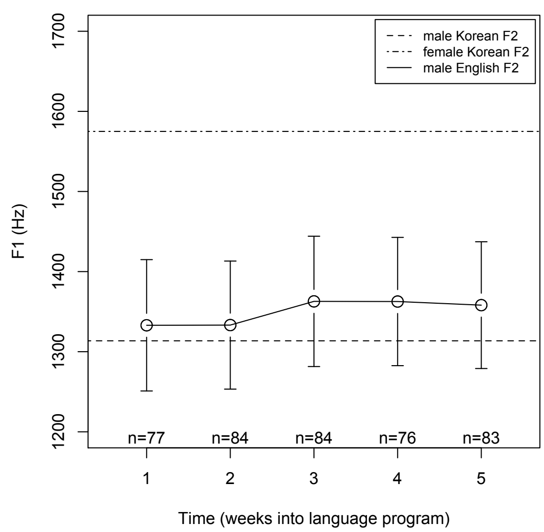
(a) Mean F_1, females



(b) Mean F_2, females



(c) Mean F_1, males



(d) Mean F_2, males

Figure 5.18: Mean F_1 and mean F_2 of the English vowel space (solid lines) over time in comparison to the Korean vowel space (dotted lines), by formant and talker gender: (a) mean F_1, females; (b) mean F_2, females; (c) mean F_1, males (in comparison to males and females); (d) mean F_2, males (in comparison to males and females). Error bars indicate 95% confidence intervals.

of the male English vowels closer to the higher mean F_2 of the female Korean vowels (Figure 5.18d). In both F_1 and F_2, however, the male English vowels stay far below the formant values of female Korean vowels.

In sum, female learners show a decreasing developmental pattern in the mean F_1 of their L1 English vowels that approximates the mean F_1 of their L2 Korean vowel targets, whereas they show little change in mean F_2. On the other hand, male learners show small increases in both mean F_1 and F_2 of their L1 English vowels, which are also most likely approximating female L2 Korean vowel targets.

5.4 Discussion

This case study of vowels provided evidence supporting three hypotheses of Chapter 2: early L1 phonetic drift in L2 acquisition, assimilatory L1 phonetic drift in late L2 learners, and L1 phonetic drift of vowels at a systemic level. Consistent with the predictions of Section 5.1.2, phonetic drift of English vowels resulted in an upward shift in the F_1 dimension, though no significant shift in the F_2 dimension. Moreover, although there was inter-speaker variability in vowel production and change, no clear dialect-related differences in drift emerged (although this may have been because the task was not successful at eliciting dialectal pronunciations). In this section, we review possible explanations for some asymmetries in vocalic drift, including the non-occurrence of the predicted drift in F_2, consider alternative explanations for the results, and compare the findings to those of other relevant studies in the L2 speech and speech accommodation literature.

The current results showed an assimilatory kind of phonetic drift in novice L2 learners' L1 vowels that approximated L2 vowels at a global level. However, a notable asymmetry was observed between female and male learners. While the L1 vowel space of female learners drifted towards the L2 vowel targets of female native speakers, resulting in an overall raising of the female L1 vowel space, the L1 vowel space of male learners did not drift towards the L2 vowel targets of male native speakers, which would have resulted in a similar raising of the male L1 vowel space. Instead, males' L1 vowels also drifted towards female L2 vowel targets, presumably due to the fact that their L2 input came predominantly from female native speakers of L2. Consequently, the pattern of phonetic drift in the male L1 vowel space—albeit not statistically significant with only three male learners—is almost the mirror image of that found in the female L1 vowel space: the male L1 vowel space lowers and, to a certain degree, advances, rather than raising. This contrast is visualized in Figures 5.19 and 5.20, which plot the change in the L1 English vowel space of female learners and male learners, respectively, in comparison to the Korean vowel space of female native speakers.

Before these findings are discussed further, an alternative explanation of the results should be considered—namely, that the spectral changes in English vowels are

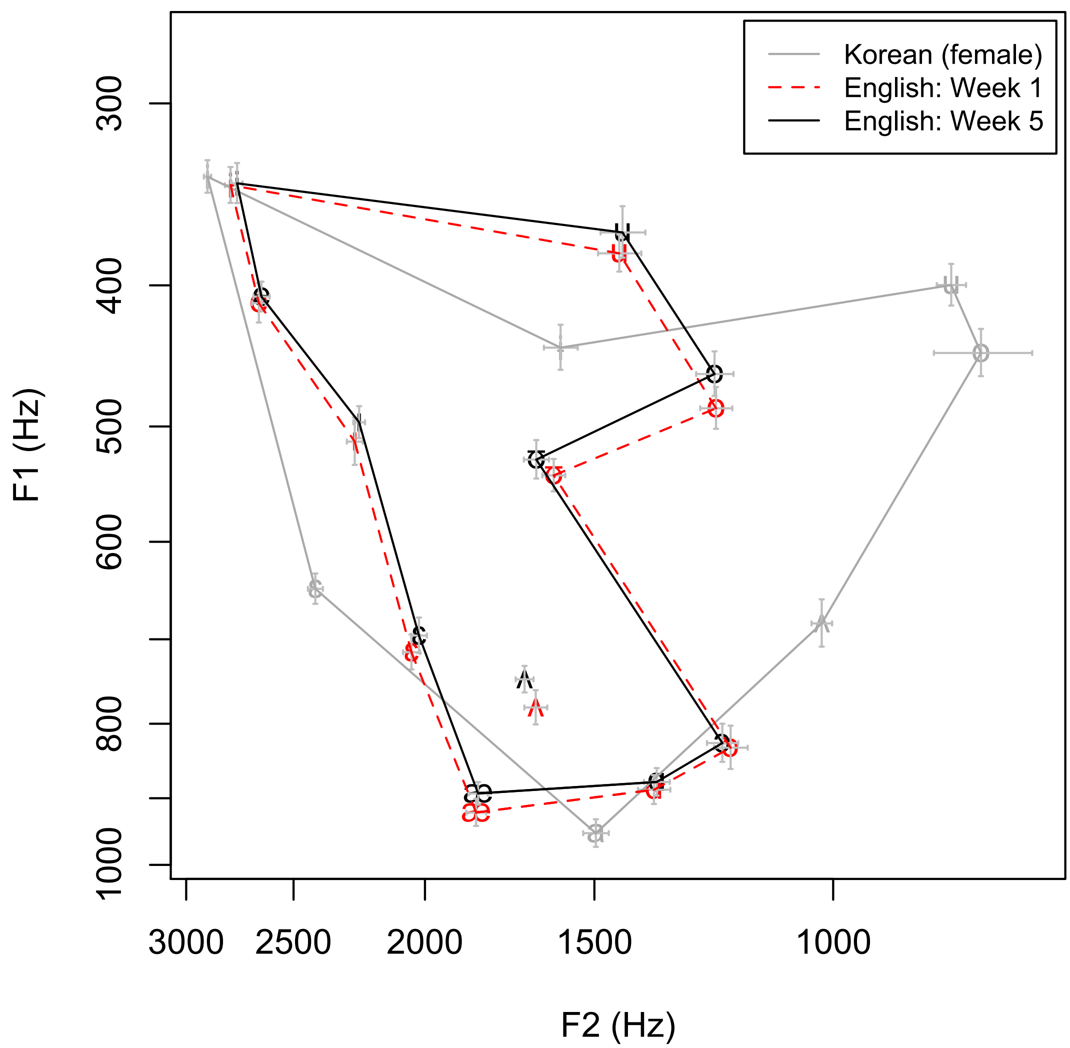


Figure 5.19: Phonetic drift of the English vowel space, female L2 learners of Korean. The scale of both axes is logarithmic. Plot symbols are the standard IPA transcriptions of the vowels. Means for Korean female vowels are plotted in gray (solid lines), for English vowels in Week 1 in red (dotted lines), and for English vowels in Week 5 in black (solid lines). Error bars indicate ±1 standard error about the mean.

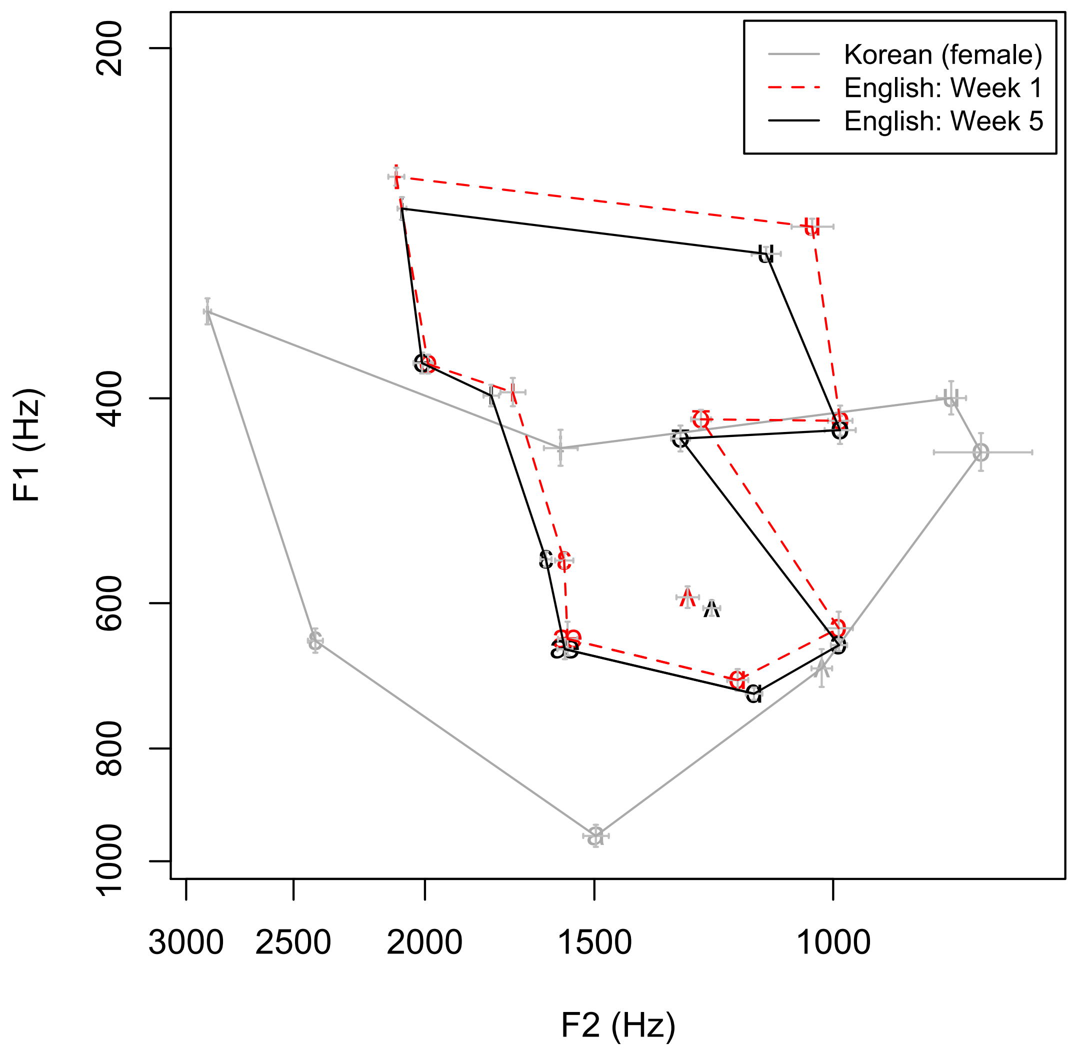


Figure 5.20: Phonetic drift of the English vowel space, male L2 learners of Korean. The scale of both axes is logarithmic. Plot symbols are the standard IPA transcriptions of the vowels. Means for Korean female vowels are plotted in gray (solid lines), for English vowels in Week 1 in red (dotted lines), and for English vowels in Week 5 in black (solid lines). Error bars indicate ±1 standard error about the mean.

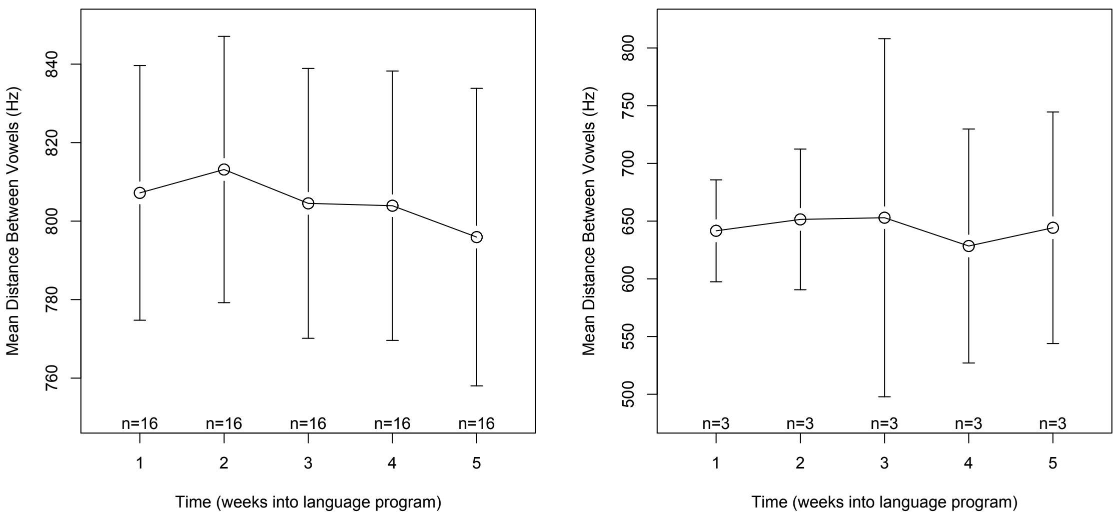
epiphenomenal, an artifact of changes in vowel duration. Perhaps females' F_1, for example, was found to generally decrease because their speech rate in the experiment increased over time, resulting in shorter vowels and more "target undershoot" (Lindblom 1963) in later weeks. While superficially reasonable, this account is flawed in two ways. First, it is inconsistent with duration data, which show that English vowels were produced, on average, as slightly longer in Week 5 than in Week 1 (approximately 10 ms longer for female learners and 2 ms longer for male learners). Second, it is unclear under such an account whether overall there should be any net change in mean formants of the vowel space as a whole, since the centralizing effect of target undershoot in fast speech should affect both low vowels and high vowels; in other words, one would expect the high vowels to increase in F_1 at the same time the low vowels decrease in F_1. Yet there is no such complementarity, as seen in Figure 5.19. The opposite argument for "target overshoot" (i.e., more exaggerated productions over time) based upon the slightly longer vowel durations in Week 5 is equally implausible, since in this case one would again expect different effects on the high and low vowels—a decrease in F_1 for the high vowels, but an increase in F_1 for the low vowels. The fact that phonetic drift goes in the same direction for vowels of different types suggests that these developments are not attributable to changes in vowel duration and, moreover, that they are not specific to individual vowels.

It remains a question why phonetic drift in L1 vowels seems to occur at a macro level—the level of the entire vowel space—rather than at a micro level targeting individual vowels. One explanation is that systemic pressures towards maximum vowel dispersion (Liljencrants and Lindblom 1972; Crothers 1978; Maddieson 1984; Lindblom 1986; Schwartz, Boë, Vallée, and Abry 1997; Johnson 2000) continue to exert their influence while L1 vowels are drifting. Thus, a decrease in the F_1 of a particular low vowel, for instance, could have a "domino effect", inducing similar spectral changes in other vowels so as to maintain the same degree of vowel dispersion. To put it another way, drift in one vowel may upset the system, and consequently other vowels may drift in order to return the system to its previous state of equilibrium. Such interconnectedness of vowels is well-documented in the L2 speech literature. L1 Portuguese learners of English, for instance, have been found to experience a decline in their production of English /ɛ/ at the same time their production of English /æ/ and overall accent improves (Major 1987b), and L1 Dutch learners of English have been found to experience a similar decline in production of English /u/ contemporaneous with improvements in their overall proficiency and, by implication, improvements in production of other vowels (Flege 1992). These studies thus suggest that the most informative view of a vowel's production is one that takes in account not just one vowel, but the relationship of that vowel to other vowels within the vowel system.

As discussed in Chapter 2 (Section 2.4.3), Guion (2003) provided a similar account of phonetic drift in the L1 Quichua vowels of L1 Quichua-L2 Spanish bilinguals. The raising of the three Quichua vowels /ɪ, a, ʊ/ was explained in terms of a systemic development motivated by pressures toward "sufficient" vowel dispersion between distinct vowel categories with minimal movement. It was observed in Chapter 2, however, that the apparently coordinated raising of the three Quichua vowels could theoretically have just been the sum total of three separate vowel-to-vowel effects that happened to go in the same direction. The present study, which looked at a more crowded L1 vowel inventory and again found a system-level shift in the same direction, supports Guion's original conclusion that the raising of the Quichua vowels happened due to a system-level movement. In contrast to the case of Quichua, however, the raising observed in the present study could not have been motivated by maximization of vowel dispersion in a shared L1-L2 phonetic space, since the results show that the phonetic drift did not in fact produce increased dispersion of English and Korean vowels. When the mean acoustic distance between English and Korean vowels (i.e., the average of the distances between every possible English-Korean vowel pair) was calculated for each time point, it was found that this index of cross-linguistic vowel spacing did not increase over time for either female or male learners (Figure 5.21). On the contrary, for female learners the mean spacing between vowels *decreased* over time, a trend that was marginally significant [$F(4, 60) = 2.24, p < 0.1$]. The system-level phonetic drift found in this study is thus not amenable to an explanation in terms of dispersion maximization.

Far from dissimilation between vowels for the sake of maximizing dispersion, the proposal of this study is that the systemic drift of L1 vowels found here is ultimately an instance of assimilation to the L2 vowel system. The discrepancy with the findings of Guion (2003), then, is likely due to differences in the age of onset of L2 experience between the two studies: the participants in her study who manifested dissimilatory drift were relatively early bilinguals (thus increasing the likelihood of cross-language dissimilation, according to the SLM), whereas all of the participants in the current study were late adult L2 learners. As for the basis of assimilatory phonetic drift at the level of the vowel system, in Section 5.1.2 it was suggested that this sort of shift might occur via cross-language links between L1 and L2 vowel systems at a global level (e.g., overall F_1 and F_2 levels); however, the nature of the structure linked across languages in this way is not yet clear. Perhaps L2 learners, tracking a long-term average spectrum of L2 in comparison to L1, link the F_1 of the L2 spectrum, for example, with that of the L1 spectrum, and in this way production of L1 shifts in the direction of L2 at the level of overall F_1. Alternatively, the comparison between grand spectral means may occur at a higher level—for instance, averaging over vowel types (as seen in Table 5.4) rather than tokens. Further work on L1s and L2s with vowel inventories that have different frequency characteristics is necessary to distinguish these two accounts.

Above it was observed that there was an asymmetry between female and male learners in terms of phonetic drift, which occurred in opposite directions for these two groups. An additional asymmetry is evident in the spectral properties that drift, in particular with respect to the conditions under which they do so. In Figure 5.18 it was seen that female learners' F_1 in L1 vowels decreased in approximation to the lower F_1



(a) Mean vowel spacing, females

(b) Mean vowel spacing, males

Figure 5.21: Mean spacing between model (native female) Korean vowels and English vowels over time, by talker gender: (a) female L2 learners; (b) male L2 learners. Error bars indicate 95% confidence intervals.

norms of female L2 vowels, and that male learners' F_1 and F_2 in L1 vowels increased in approximation to the higher F_1 and F_2 norms of female L2 vowels. However, female learners' F_2 in L1 vowels did not decrease in approximation to the lower F_2 norms of female L2 vowels. Puzzling as this finding is, it recurs throughout the literature on inter-language phonology. In Flege (1987b), for example, native French-speaking L2 learners of English manifested phonetic drift of their L1 French /u/ in the direction of L2: F_2 of French /u/ was found to increase in approximation to the higher F_2 of English /u/. However, the English /u/ of native English-speaking L2 learners of French did not decrease in F_2 to become more similar to French /u/. Similarly, in Chang et al. (2010), native Mandarin-speaking L2 learners of English manifested phonetic drift of their L1 Mandarin /u/, which was produced with F_2 values higher than native Mandarin norms, while native English-speaking L2 learners of Mandarin did not manifest phonetic drift of their L1 English /u/, which was generally produced in the range of phonetic norms for American English.

While the reason for this asymmetry in vocalic drift is not entirely clear, one possibility is that it is based in fundamental properties of human audition. Specifically, due to the structure of the basilar membrane in the cochlea, in which the larger, more compliant apical area responds selectively to lower frequencies over higher frequencies (von Békésy 1960; Goldstein 2010), humans are relatively more sensitive to differences between low frequencies than to differences between high frequencies. It follows that differences in F_1 may be generally more perceptible than equivalent differences in F_2 because of their respective locations in the frequency range. This perceptual bias may account for why phonetic drift in female learners' L1 vowels occurred in F_1, but not in F_2. It may also account for why the lower F_2 of male learners' L1 vowels appeared to drift towards the F_2 of female L2 vowels, while the higher F_2 of female learners' L1 vowels did not. However, it cannot account for why low-F_2 vowels of L1 generally seem to drift more readily towards corresponding high-F_2 vowels of L2 than high-F_2 vowels of L1 drift towards corresponding low-F_2 vowels of L2. A different perceptual bias seems to be responsible for this asymmetry—namely, a directional one. A directional bias exists in frequency discrimination, such that changes in frequency are more accurately perceived in the right ear (the ear for which a linguistic advantage has been found) when the frequency change is ascending as opposed to descending (Murray 1980). Such a perceptual bias may underlie the asymmetry described above, in that a higher F_2 in L2 may be more likely to be perceived as different from F_2 in L1 than is a lower F_2 in L2. If this were the case, it would follow that a higher F_2 in L2 would more often induce L1 phonetic drift.

Though the pattern of phonetic drift in stops discussed in Chapter 4 bore some strong resemblances to the pattern of phonetic imitation in stops, there are significant differences between the pattern of phonetic drift in L1 vowels seen here and the pattern of phonetic imitation in L1 vowels seen in Babel (2009b). Whereas in Babel (2009b) change in L1 vowel production (in approximation to vowels of an L1 model talker) was localized mainly to the F_1 of the low vowels /æ, ɑ/, in the present study change in L1 vowel production (in approximation to vowels of an L2) occurred at a global level, affecting the vowel space as a whole rather than a few vowels in particular. That this change is systemic, rather than simply the sum total of assimilatory changes in individual L1 vowels, is clear from a close examination of the drift patterns of L1 vowels in relation to L2 vowels. In Figure 5.19, for example, it can be seen that while the raising of English /ʊ/ brings it closer to the nearby Korean /ɨ/, the raising of English /u/ takes it farther away from both Korean /ɨ/ and /u/, the two closest L2 vowels. The existence of many such contrasts indicates that the observed vowel shifts cannot be accounted for coherently in terms of vowel-to-vowel effects.

One obvious difference between these two studies that might be appealed to in order to account for the disparate results is the role played by lexical overlap in phonetic imitation (and not in phonetic drift). It is unclear, however, how this factor can distinguish between phonetic drift and phonetic imitation in the necessary way. In phonetic imitation, lexical overlap between the L1 input from the model talker and the L1 imitative output straightforwardly delineates a set of tokens to use in imitation—namely, tokens of the specific input word that are already in the imitator's repertoire. As argued by Babel, if imitation occurs within the bounds of such word-specific phonetics (Pierrehumbert 2002), it follows that there will be more imitation of vowels that are, within the same lexical/phonological context, more variable in their production, since there is a wider range of tokens to draw upon in matching the input. Lexical overlap cannot play this role in the case of phonetic drift, given that different languages typically have few words in common. However, this actually predicts that, if anything, there will be an even wider range of L1 vowel tokens to draw upon in approximating L2 vowels in phonetic drift, since the range may not be limited by lexical factors. Thus, the presence vs. absence of lexical overlap cannot be responsible for the differences between the L1 vowel shifts at issue.

Instead, the different results of the two studies are likely attributable to the differences in time scale involved. In shadowing tasks such as those used in Babel (2009b), L1 productions are elicited immediately following brief exposure to L1 input, while in the present study L1 productions were elicited well after prolonged exposure to L2 input, which mostly occurred in the context of classes that ended for the week at least several hours before the experiment began. Thus, by allowing for more time to synthesize and consolidate the relevant input, the task in the current study is probably more conducive to eliciting L1 vowel shifts that are influenced by global linkages to the input, rather than by vowel- or word-specific linkages. It is an interesting question, then, whether one might find similarly systemic shifts in a shadowing task if, for example, subjects were exposed to an L1 model talker for a longer period of time, or if there were a longer delay between exposure to the model talker and elicitation of productions.

On a final note, it should be observed that although the motivation for L1 vocalic drift differed between this study and the study of Guion (2003), both cases of drift happened to result in the same direction of shift—that is, vowel raising. As such, it would be useful to extend this work combining the type of L1-L2 pairing investigated by Guion (2003) with the population investigated in the current study (and vice versa) in order to determine whether the observed patterns of L1 vowel raising do in fact derive from the principles that have been discussed in this dissertation, as opposed to some universal tendency for vowels to be raised in contact situations. If, for example, the present results are indeed the product of adult learners' assimilation of L1 to L2, then we expect to find the complementary result—that is, vowel lowering—looking at adult L2 learners whose L1 has a relatively low overall F_1 level (e.g., Korean) and whose L2 has a relatively high overall F_1 level (e.g., English). Similarly, it is reasonable to predict that early bilinguals with the reverse kind of vowel inventory pairing (e.g., early L1 English-L2 Korean bilinguals) would also exhibit vocalic drift of a lowering variety. Controlled studies of L1 vocalic drift in these types of linguistic situations would greatly broaden our understanding of the basis of this phenomenon and the ways in which it is constrained by the structural, auditory, and perceptual biases highlighted in this chapter.

Chapter 6

General Discussion

6.1 Introduction

This chapter attempts to synthesize the present findings with previous reports of L2 effects on L1 in service of a more well-defined notion of phonetic drift. In addition, it discusses the relevance of the findings for several aspects of linguistic theory, including the basis of cross-linguistic similarity, the place of different types of experience in an exemplar model of phonology, and the path of contact-induced sound change. The chapter concludes by considering the implications of the results for language learners, language teachers, and researchers in the social sciences.

6.2 Redefining Phonetic Drift

In Chapter 1, phonetic drift was defined as "phonological restructuring in the first language during second language acquisition". Chapter 2 focused on the result of this restructuring—namely, "subtle shifts in L1 sounds resulting from experience with similar, but non-identical L2 sounds". In this section, a unified account of these shifts is presented in an exemplar framework (Hintzman 1986; Johnson 1997, 2006; Goldinger 1998; Pierrehumbert 2001, 2002), including both their cause and their developmental path.

6.2.1 Basis of Phonetic Drift

In order for there to be detectable, assimilatory phonetic drift in L1 during L2 acquisition, specific conditions need to obtain. Assimilatory drift arises only when an L2 sound exists at what we might think of as a "sweet spot" of cross-linguistic distance from the closest L1 sound. According to the SLM, if the L2 sound is very different from the L1 sound, it will be perceived as a "new" sound and thus will not be linked to the phonetic category for the L1 sound to begin with. On the other hand, if the L2

Figure 6.1: Continuum of similarity of L2 sounds to L1 sounds. "New" L2 sounds are the least similar to L1 sounds; "identical" L2 sounds, the most similar; and "similar" L2 sounds, intermediate in similarity.

sound is hardly different from the L1 sound, it may constitute an "identical" sound, such that its perceptual linkage to the L1 sound results in no noticeable change in L1 production (i.e., there is only "positive transfer"). The type of L2 sound that causes noticeable assimilatory drift is instead a "similar" sound—an L2 sound that resembles an L1 sound to such a degree that it undergoes equivalence classification with the L1 sound, yet is still different enough that there is significant cross-linguistic perceptual distance between the two sounds. Given the gradient nature of similarity, "identical", "similar", and "new" L2 sounds might be thought of as ranges in a continuum of cross-linguistic similarity to L1, as shown in Figure 6.1. At the end of minimal similarity to L1 are "new" L2 sounds, while at the end of maximal similarity to L1 are "identical" L2 sounds. "Similar" L2 sounds occupy a broad swath in the middle of the continuum, spanning a considerable range of cross-linguistic similarity. As such, some "similar" sounds lying near the IDENTICAL end of the continuum resemble L1 sounds closely, while others lying near the NEW end of the continuum resemble L1 sounds only marginally.

Crucially, there must be sufficient dissimilarity between similar L1 and L2 sounds in order for the L1 sound to manifest phonetic drift. In this regard, it is important to note that even when an L2 sound is identified with an L1 category, learners continue to have "access at an auditory level to cross-language phonetic differences" (Flege 1995:258). However, for phonetic drift to occur the differences must be significant; otherwise, the distinction between the L1 and L2 sounds may be too small to be perceived. If cross-linguistic similarity in just one dimension is considered, it is assumed that the boundary between "similar" L2 sounds and "identical" L2 sounds occurs at a distance from the L1 sound of approximately one just-noticeable difference (JND) in the relevant phonetic dimension. An L2 sound that lies more than one JND away from an L1 sound can be distinguished from it, but an L2 sound that lies less than one JND away is not perceived as any different and is thus treated as the same sound. This accords with the view of Pierrehumbert (2001) that speech tokens that are too similar to be discriminated are stored as if they were identical:

> ...it is reasonable to suppose that speech tokens differing by less than one JND in f0 are stored as if they had identical f0s. Similar constraints on the resolution of all other perceptual dimensions would motivate granulariza-

> tion of the phonetic parameter space as a whole. As a result, an individual exemplar—which is a detailed perceptual memory—does not correspond to a single perceptual experience, but rather to an equivalence class of perceptual experiences. (Pierrehumbert 2001:141)

Thus, it is claimed that phonetic drift is dependent on an L1 sound being perceptually distinguished from a close L2 sound. If no difference between the two sounds is perceived, phonetic drift does not occur because tokens of the L2 sound are not stored uniquely. Without any separate encoding, the L2 sound does not have an independent existence, so there can be no interference from it in L1.

Given a similar, yet distinguishable L2 sound, phonetic drift in L1 arises as a product of the way in which the exemplar clouds of L1 and L2 categories overlap with each other. Production of an L1 sound shifts when exemplars of the L2 sound—especially those that lie far from the central tendency of the L1 sound—are incorporated into the exemplar cloud of the L1 sound as a result of equivalence classification. The incorporation of phonetically peripheral L2 tokens into the L1 cloud changes the average of the L1 exemplar distribution and, thus, the properties of the most likely production of the L1 sound. Whether phonetic drift is detectable will depend on whether the difference in category distributions is large enough to make a statistically significant dent in the distribution of the L1 sound. More generally, the magnitude of the drift is expected to increase with the magnitude of cross-linguistic dissimilarity. To put it another way, the wider the gap between similar L1 and L2 sounds, the more room there will be for the L1 sound to drift.

Following from the assumption that "similar" L2 sounds lie at a distance of at least one JND from L1 sounds, the minimal amount of (dis)similarity between L1 and L2 required for phonetic drift to occur is posited to be one JND. Indeed, the results of this study are consistent with a criterion of one JND, in light of the size of JNDs that have been reported in the literature for duration (Lehiste 1970; Huggins 1972; Klatt 1976), including VOT (Hazan et al. 2009), and for frequencies in the range of f_0, F_1, and F_2 (Roederer 1973; Harrison 1996; Oglesbee and Kewley-Port 2009). JNDs for VOT in quiet listening conditions are cited at 16–23 ms, and these figures are consistent with the patterns of VOT drift found in Chapter 4. No VOT drift was found in the English voiced stops, which differed in terms of VOT norms from the perceptually similar Korean fortis stops by less than 16 ms (mean of 4 ms; range of 1–7 ms). In contrast, significant VOT drift was found in the English voiceless stops, which differed in terms of VOT norms from the perceptually similar Korean aspirated stops by just about the right amount for the difference to be noticed (mean of 23 ms; range of 12–30 ms). JNDs for f_0 are cited at 4–5 Hz, and these figures are also consistent with the patterns of f_0 drift found in Chapter 4. Phonetic drift in f_0 onset was observed in the English voiced and voiceless stops, which probably differed in terms of f_0 norms from similar Korean stop series by well over 5 Hz in both cases. In the case of vowel formants, JNDs for F_1 are cited at 19–45 Hz (19–35

Hz for front vowels and 36–45 Hz for non-front vowels), while JNDs for F_2 are cited at 76–161 Hz (141–161 Hz for front vowels and 76–80 Hz for non-front vowels). These figures are partly consistent with the patterns of vocalic drift found in Chapter 5. The increases in male learners' F_1 and F_2 for English are as expected, since the mean F_1 and F_2 for (female) Korean were, respectively, 91 Hz and 242 Hz higher in Week 1, differences well beyond the JNDs for F_1 and F_2. The decrease in female learners' F_1 for English is as expected, too, since the mean F_1 for Korean was 33 Hz lower in Week 1, a difference that was also probably able to be perceived. On the other hand, no significant drift occurred in female learners' F_2 for English, even though the difference between the mean F_2 for Korean and the mean F_2 for English—at 160 Hz in Week 1—should have been perceptible according to the above numbers.

There are two possible explanations for the lack of drift in female learners' F_2. One account is that drift in females' F_2 was blocked by the reduced vowel spacing resulting from the assimilatory drift in females' F_1. Whereas the phonetic drift of male learners' English vowels did not significantly affect the degree of vowel spacing for them (Figure 5.21b), as their English vowels were relatively distant from female Korean vowels, the phonetic drift of female learners' F_1 for English vowels resulted in tighter vowel spacing (Figure 5.21a), as their English vowels were relatively close to female Korean vowels. Consequently, the pressure to maintain cross-linguistic contrast may have headed off additional drift that would have further decreased the already reduced degree of dispersion in the vowel system. Another contributing factor could have been directional asymmetries in perception. As mentioned in Chapter 5, the lack of drift in females' F_2 for English may have been due in part to a directional bias in frequency discrimination. Such a bias may also help to account for asymmetries in vocalic drift that have been observed in other studies (e.g., Flege 1987b; Chang et al. 2010). The generality of such asymmetries and their existence in other domains of contrast are issues that should continue to be investigated in future work.

The role of somatosensory feedback in shaping patterns of phonetic drift should also be considered. In Chapter 2, it was noted that studies of speech adaptation and speech accommodation have produced similar results showing that L1 speech production is highly influenced by external input. Another way in which these research areas have converged is in the finding that change in L1 production in response to an environmental stimulus is typically incomplete. Talkers' production shifts in approximation to a stimulus (or in compensation for a distortion), but usually does not shift all the way. The incompleteness of such production shifts has been attributed to the grounding influence of somatosensory feedback. Somatosensory feedback may also play an important role in moderating the magnitude of L1 phonetic drift, in that talkers may be prevented from shifting the production of an L1 sound to match an L2 sound by their sense of what production of the L1 sound is supposed to feel like. Together with the general pressure to maintain cross-linguistic contrast, the influence of somatosensory feedback renders it highly unlikely for phonetic drift to result in the merger of L1 sounds with L2 sounds—a pattern that is, as of yet, unattested in late L2 learners. However, somatosensory feedback may have grounding effects on L1 production patterns to varying degrees in different phonetic dimensions. It is reasonable to suppose, for example, that the effect of somatosensory feedback might be stronger for contrasts based in articulatory targets and postures than for contrasts based in articulatory timing and prosodic features. Whether this is actually the case is an empirical question.

Although the three cases of phonetic drift found in this study were all assimilatory to L2, they differed in the nature of the cross-language linkage that provided the basis for the shift in L1 production. In contrast to the segmental focus of the L2 speech models discussed in Chapter 2, phonetic drift in this study was observed on both a segmental level and a non-segmental level. In the case of VOT, the VOT of L1 English voiceless stops drifted towards the longer VOT of L2 Korean aspirated stops. However, differences in the magnitude of drift across places of articulation suggested that the cross-language linkage in this case was not just between the natural class of L1 voiceless stops and the natural class of L2 aspirated stops, but also between L1 /p/ and L2 /p^{h}/, L1 /t/ and L2 /t^{h}/, and L1 /k/ and L2 /k^{h}/. Thus, phonetic drift in VOT appears to draw upon cross-language linkages established at the subphonemic (natural class) level as well as the phonemic (segmental) level. On the other hand, phonetic drift in frequency components was found to be influenced by cross-language linkages on a global level. Drift in f_0 onset was not limited to the English voiced and voiceless stops identified with Korean fortis and aspirated stops. Instead, the higher overall f_0 of Korean (resulting from the elevation of f_0 following laryngeally marked onsets such as fortis and aspirated stops) led to the elevation of f_0 in nearly all English words, including the ones without any sort of consonantal onset, although drift in f_0 was still greater for stop-initial English words than for vowel-initial English words. Phonetic drift in f_0, therefore, cannot be based solely on cross-language linkages between similar natural classes of stops. On the contrary, the findings of work on speech adjustments in cochlear implant users suggest that suprasegmental, or "postural", properties of speech such as f_0 may be modulated by a control mechanism different from the one responsible for the control of segment-level articulation. This control mechanism, which also appears to play a role in the modulation of parameters such as speech rate, sound level, and airflow, is likely to be non-language-specific in whole or in part. It follows that experience with f_0 in L2 could influence f_0 production in L1 on a global level since f_0 in both L1 and L2 would be maintained by the same general control structure. With regard to vowel formants, similar to the case of f_0 phonetic drift occurred on a global level, with a general shift in F_1 that could not be accounted for coherently in terms of individual vowel shifts. Thus, the cross-language linkage underlying this drift cannot be based on a segmental unit of analysis; instead, it is probably established over the entire vowel system, as suggested in Chapter 5. Taken together, the findings of this study provide evidence of a cross-language phonetic phenomenon that is much more multifaceted than the segmental interference constituting the focus of current L2 speech models.

Consistent with the predictions laid out in Chapter 2, phonetic drift in the adult L2 learners examined here was found to be assimilatory, yet cases of dissimilatory drift have also been reported, especially for other populations. A complete account of phonetic drift must address the question of how to predict when drift will be dissimilatory versus assimilatory. According to the SLM, dissimilation occurs when an L2 sound, instead of being linked to an L1 phonetic category, precipitates the formation of its own phonetic category. Given adjacent L1 and L2 categories in a shared phonetic space, dissimilation between them serves to enhance cross-linguistic contrast. However, it is not well-understood how far from L1 an L2 sound needs to be both to be perceived as a "new" sound (as opposed to a "similar" sound) and to cause dissimilation from an adjacent L1 phonetic category. These two events—classification as a "new" sound and dissimilation from an L1 category—impose opposite demands in this respect. On the one hand, a high degree of disparity from L1 is necessary for an L2 sound to be classified as a "new" sound meriting its own phonetic category. However, at the same time a substantial degree of proximity to L1 is necessary for dissimilation to operate, as there is probably no need to dissimilate categories that are already very far apart. As such, it may not be the case that L2 category formation necessarily results in dissimilatory drift from L1, as noted by Flege (2002:239), who acknowledged "the possibility of an L2 sound that is so distant in the phonetic space from the closest L1 category that a category established for it would not influence realizations of the Ll category". Thus, just as there seems to be a "sweet spot" of distance from L1 for inducing assimilatory phonetic drift, so too might there be a "sweet spot" of distance from L1 for inducing dissimilatory phonetic drift: an L2 sound has to be distant enough from L1 sounds to be considered a "new" sound, yet not so distant that there is no L1 sound within its vicinity. What remains unclear is if there is indeed a degree of distance from L1 that qualifies as too far to cause dissimilation. Moreover, the SLM's formulation of the basis for cross-language dissimilation in terms of segmental categories implies that dissimilation can only occur at the segmental level, but whether this holds true remains to be seen.

6.2.2 Time Course of Phonetic Drift

The immediate nature of perceptual linkage between similar L1 and L2 sounds, which were conceptualized as a complex of overlapping exemplar clouds, implies that L1 phonetic drift should occur from the onset of L2 exposure. The hypothesis of early phonetic drift was supported by the findings of this study, which showed significant L1 production shifts in VOT, as well as in f_0 and F_1, weeks into a formal program of L2 learning. However, drift was found to occur more rapidly in VOT and f_0 than in F_1. This difference suggests that the nature of the cross-language linkage giving rise to phonetic drift, as well as the degree to which the specific phonetic dimension at issue is affected by somatosensory feedback, may play a role in determining the exact trajectory of L1 production changes. To be specific, it may be that vowel systems, by virtue of their size, take longer to drift due to greater inertia relative to segment-sized units. Moreover, vowel qualities, as categories differing in terms of tongue posture specifically, may be relatively more strongly prevented by somatosensory feedback from drifting away from monolingual L1 phonetic norms. Thus, it seems likely that the time course of phonetic drift will vary depending on a number of considerations.

In Chapter 2, it was claimed that the design of this study provided the conditions least likely to produce evidence of L1 phonetic drift: adult monolingual L1 speakers, an L2 unrelated to L1, and formal L1 speech. In reality, one aspect of the design actually encouraged phonetic drift—production of L2. As described in Chapter 3, study participants produced L1 shortly after completing a different set of production tasks in L2. Having produced L2 beforehand is expected to have affected performance in the L1 production task because it would have activated L2 representations that could have interfered with L1 production. The reason why this does not negate the finding of phonetic drift is because the design remained the same, giving L2 the same opportunity to influence L1 at all time points. Thus, the fact that the influence of L2 on L1 increased over time suggests that acquiring L2 experience altered the phonetic representations shared between L1 and L2 such that L2 was able to have a greater effect on L1. In terms of an exemplar model, it is proposed that what the accrual of L2 experience accomplishes is heavier weighting of the L2 component of a representation via an increase in the relative mass of L2 exemplars (i.e., "strength in numbers"). L2 exemplars that fall at the periphery of the phonetic space of an L1 sound can exert a greater effect on the central tendency of the L1 exemplar cloud the more of them there are. In this way, production of the L1 sound may drift farther away from monolingual norms as more tokens of a close, but only partly overlapping L2 sound are stored and associated with the L1 category.

The net effect of L2 exemplars on an L1 category depends not only on their numbers, but also on their relative strength, which is tied to the recency of their experience. Recall that in their study of the effects of recent linguistic experience on voiceless stop production, Sancier and Fowler (1997) found evidence of slight, but significant drift in VOT towards the VOT norms of the most recently experienced language. After months of immersion in English (in which long-lag VOT is typical for voiceless stops), their L1 Portuguese-L2 English speaker produced longer VOTs than after months of immersion in Portuguese. The magnitude of the drift was small, however—on the order of 5 ms. The relatively small size of this drift was attributed to the dueling influences of less numerous, more recently experienced L2 exemplars of voiceless stops versus more numerous, less recently experienced L1 exemplars of voiceless stops. As the authors state, "if these past experiences perceiving and producing a voiceless stop also affect production, but individual recent experiences are relatively more potent than individual distant past experiences, we can explain both why recent experience has a measurable effect at all and why the effect is so small" (Sancier and Fowler 1997:432). In other words, recent L2 experience is able to make a dent in production of an L1 category in spite of the greater mass of L1 exemplars because the recency of the L2 experience boosts the strength of the L2 exemplars. Nonetheless, the implication of Sancier and Fowler (1997) is that recency can only do so much, as the mass of L1 exemplars remains much greater than the mass of L2 exemplars for a late L2 learner. In the case of their speaker, for instance, L1 had essentially a fifteen-year head start on L2 in accrual of experience.

In comparison to Sancier and Fowler (1997), the present study found a much stronger effect of recent L2 experience on L1. The results of Chapter 4 showed that weeks, not months, of experience were sufficient to induce VOT drift in voiceless stops on the order of 20 ms. This disparity with the results of Sancier and Fowler (1997) may be due to the more recent and more intense nature of the L2 experience in the present study, in which participants had spoken in L2 continuously for several minutes before they began the L1 production task. By requiring active use of linguistic representations, speaking in L2 may activate L2 exemplars in a more robust way than just listening to L2 can, and more robust L2 activation could allow for a more pronounced effect of L2 on L1 production. Though Sancier and Fowler (1997) measured their speaker's production in the U.S. right before she departed for Brazil and right after she returned, it is not clear how much she had been speaking (as opposed to just hearing) L1 and L2 immediately preceding her participation in the production task; presumably, she had mostly been listening to instructions. It is possible that the authors would find a greater amount of drift if they measured the speaker's VOT in L1 right after she had an extended conversation in L2, for example. On the other hand, it is also possible that this speaker, as a late bilingual with upwards of twelve years of L2 experience, is simply not comparable to the novice L2 learners examined in the present study. Extensive and current use of both L1 and L2 might have made her more adept at switching between the two languages, thus minimizing the degree of short-term L2 interference in L1. In addition, the effect of somatosensory feedback may have been greater in the Portuguese-English speaker than in the English-Korean learner, since in the former case phonetic drift was potentially causing production to cross over from one phonetic category (voiceless unaspirated stops) to another (voiceless aspirated stops), while in the latter case phonetic drift was not causing production to stray outside one phonetic category (voiceless aspirated stops).

Regardless of the explanation for the different results in the two studies, having participants in the current study engage in L2 production stands to magnify the effect of accruing L2 experience on L1 production by encouraging activation of a more expansive range of L2 exemplars than in L2 perception. This is what is meant above by "more robust" activation of linguistic representations in speaking. Because production is more spontaneous than perception (which always occurs in direct response to an external stimulus), the range of exemplars activated when speakers produce an utterance is probably wider than the range of exemplars activated when speakers match an acoustic signal against their current store of exemplars. It follows that more of the additional mass of exemplars gained through recent L2 experience is likely to be activated to a high degree through production of L2. As a consequence, longitudinal increases in L2 influence on L1 production may have been seen especially clearly in this study since the preceding L2 production task could have highlighted differences across time points in the mass of L2 exemplars accrued to that point, many of which would fall in the periphery of the exemplar cloud of an L1 category and thus stand to shift its production significantly given sufficient activation. Typically these marginal L2 exemplars might have been activated only weakly in activation of the L1 category, thereby limiting their effect on its production, but it is reasonable to suppose that the L2 production task provided an activation boost that helped these exemplars reach an activation level high enough to exert the large effect found in this study.

The question to ask, then, is whether early phonetic drift would be found even without such an activation boost from L2 production. One fact about the results strongly suggests that phonetic drift would be found in this case as well—namely, its perseverance throughout Experiment 2E (the English production experiment). If the phonetic drift found in Experiment 2E was entirely a function of short-term, recent activation of L2 exemplars in Experiment 1K (the Korean production experiment), the effect would be expected to diminish over time with the decay of short-term activation. Thus, phonetic drift should be more pronounced at the beginning of Experiment 2E, when recently activated L2 exemplars would have been able to exert the greatest influence on L1 production, than at the end, after their activation would have already decayed to a considerable degree. Such a decline in L2 influence over the course of Experiment 2E is not found, however. Rather, longitudinal L1 production shifts are just as much in evidence at the end of the experiment as at the beginning, as shown in Figure 6.2 for the phonetic drift in VOT of the English voiceless stops. Here it can be seen that the increase in VOT of the English stops has not been dampened at all over the 7–8 minutes of Experiment 2E. This pattern suggests that whatever short-term enhancement of L2 influence may have resulted from the act of producing L2 in a prior experiment, long-term modifications of L1 representations were primarily responsible for the phonetic drift observed in Experiment 2E.

The results of the case studies discussed in Chapters 4–5 provide evidence that L2 experience begins effecting long-term modifications of L1 representations right away. Beyond this, the timeline of L2-influenced changes to L1 representations is not clear. However, some studies imply that the bulk of perceptual developments during L2 acquisition takes place within the first six months of L2 exposure (Aoyama, Flege, Guion, Akahane-Yamada, and Yamada 2004; Jia, Strange, Wu, Collado, and Guan 2006; Best and Tyler 2007). If this is the case, it stands to reason that most L2-influenced developments in L1 representations may also occur within this timeframe. A leveling off of phonetic drift within a relatively brief time period seems possible, as at some point the mass of L2 exemplars probably becomes "saturated", such that the continued addition of new exemplars meets with diminishing returns—that is, a smaller net effect on exemplar clouds.

Although the current findings suggest that the starting point of phonetic drift coincides with the beginning of L2 experience, they do not have anything to say about

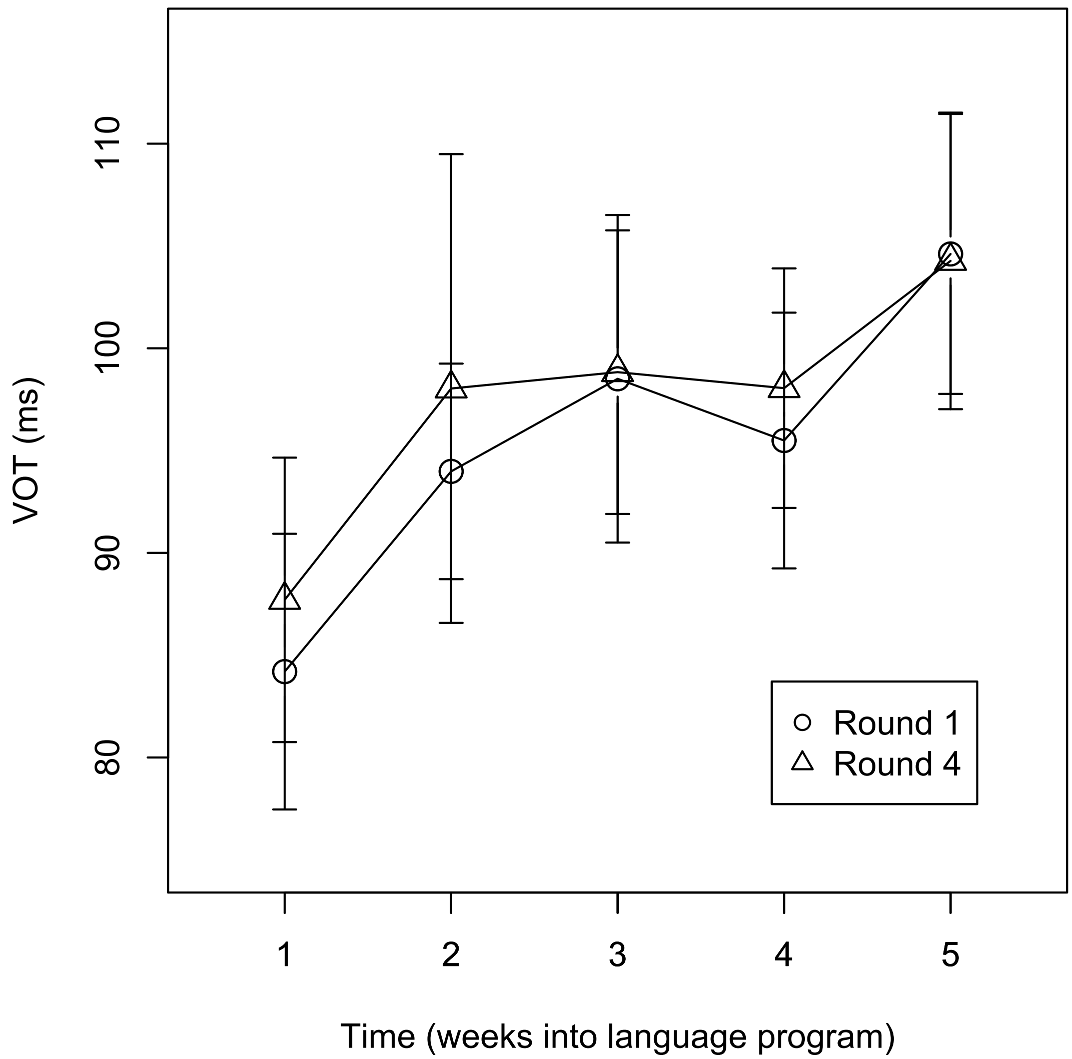


Figure 6.2: Mean VOT of English voiceless plosives over time, by token. Means for the first round of productions at the beginning of the task (Round 1) are plotted in circles, for the last round of productions at the end of the task (Round 4) in triangles. Error bars indicate 95% confidence intervals.

the endpoint of phonetic drift. One wonders whether after extensive experience in another language, an individual can ever return to producing L1 like a monolingual again. Studies on speech adaptation have shown that production adjustments made in response to altered auditory feedback persevere for a short while, but eventually dissipate after the feedback alteration is removed. Thus, it is possible that L1 phonetic drift in response to L2 exposure might also subside if L2 exposure were cut off. There are no known studies on the effects of distant, rather than recent or current, L2 experience in adulthood on L1 production, but it is probable that the degree to which L1 production drifts back to monolingual-like values in the continued absence of L2 exposure will depend greatly on the total amount of L2 experience accrued. For example, over the same period of time with only L1 exposure, less "return" drift is predicted for a speaker with twelve years of L2 experience than for one with twelve months of L2 experience. However, we know virtually nothing about the temporal progression of such "return" drift and its relation to amount of accrued L2 experience. Can an eighty-year-old immigrant with sixty years of L2 experience return to pronouncing L1 like a monolingual upon moving back to her (monolingual) native country? In the end, there may be an amount of L2 experience that is practically too great to be overridden with additional L1 experience.

6.2.3 Summary

L1 phonetic drift is based on a relationship between L1 and L2 that may be established at one of a number of levels. At the segmental level, drift occurs in an L1 sound when it becomes linked to a similar, but perceptually distinguishable L2 sound, which modifies the central tendency of the L1 exemplar cloud by virtue of its phonetic disparity from the L1 sound. At a higher level, drift occurs in a natural class of L1 sounds via class-level linkage to a parallel natural class of L2 sounds, or in an L1 system via global links to L2, which may comprise mean spectral landmarks (e.g., overall F_1) and/or shared speech control mechanisms. Assimilatory drift results in L1 production approximating L2 norms; however, complete convergence is prevented by pressure to maintain cross-linguistic contrast as well as by somatosensory feedback. Though phonetic drift tends toward assimilation with increasing age of L2 learning, it may also take the form of dissimilation, which, like assimilation, may occur only when L2 sounds exist at a relatively precise distance from L1.

The time course of L1 phonetic drift commences from the onset of L2 exposure. The trajectory of drift varies according to the phonetic dimension, which is characterized by a particular type of cross-linguistic linkage as well as a particular sensitivity to somatosensory feedback, but in general the drift is rapid, resulting in significant changes in L1 production after a few weeks of L2 experience. Phonetic drift progresses as L2 exposure leads to long-term modifications to L1 representations drawn upon in production, with the effects of L2 exposure most likely leveling off after the first few months of L2 experience. However, the perseverance of phonetic drift in the absence of continued L2 exposure is an open question, allowing for the possibility that the mark of L2 experience on L1 production will in some cases be permanent.

6.3 Implications for Theory

6.3.1 Cross-Linguistic Similarity in L2 Acquisition

The notion of cross-linguistic similarity is central to models of L2 speech perception and production, and changes in perceived cross-linguistic similarity might be one of the main factors behind the appearance of "age" effects in L2 speech acquisition (Flege and MacKay 2010). Nevertheless, the way in which an L2 learner arrives at a judgment of similarity between L1 and L2 sounds is not yet well-understood. According to Johnson and Babel (2010), phonetic similarity between two sounds has three main components: raw auditory similarity, lack of phonemic contrast, and phonological linkage by productive alternation. This definition encompasses many, but not all, of the considerations incorporated into the PAM-L2's account of cross-language relationships in L2 perception. In allowing these relationships to be established at any of three levels of representation (phonological, phonetic, and gestural), the PAM-L2 currently offers the most comprehensive view of how L1 and L2 sounds may be deemed similar. However, Chapter 2 raised the issue of conflicts between different levels: how does an L2 learner decide on the closest L1 counterpart of an L2 sound when different considerations point to different L1 sounds? This study is not the first to allude to this conundrum (see, e.g., Hammarberg 1996).

Conflicts between the phonological and phonetic levels have been especially noted in the literature, yet the question of how such conflicts are resolved has not been fully addressed. It is reasonable to think that phonological parallelism plays an important role in judgments of overall cross-linguistic similarity because it has been widely shown to have a large impact on inter-language segmental mapping in loanword adaptation (Paradis and LaCharité 1997; LaCharité and Paradis 2005; Kang 2008; Chang 2009b, in press). Work on loanword phonology has indicated that very often the closest L1 adaptation of an L2 sound is not the phonetically closest L1 sound, but rather the phonologically closest L1 sound.[1] Therefore, in regard to L2 acquisition, the phonological level may be expected to exert considerable influence in establishing cross-language linkages. However, in those cases where the phonological level conflicts with the phonetic level and/or gestural level in its indication of the closest L1 sound to an L2 sound, the PAM-L2 makes no predictions as to which level prevails.

Though the PAM-L2 is silent on this matter, several L2 studies have suggested that in these cases of conflict, the phonological level takes precedence over the pho-

[1] Phonological distance in this context is usually claimed to have a featural basis, such as the number of featural changes that need to occur to go from one sound to another (as posited in the Theory of Constraints and Repair Strategies by Paradis and LaCharité 1997).

netic level. Judgments of cross-linguistic similarity based on phonological proximity are plausible in light of findings showing that perceptual assimilation does not necessarily follow from strict phonetic proximity (Strange et al. 2004), and evidence of such phonologically-based similarity is found in the literature. In languages with voiced and voiceless stops, for example, usually the phonologically matching stop series are linked to each other even when their phonetic implementations in terms of VOT suggest that the phonetically closest categories are actually the phonologically mismatched ones. This is evidenced in the study of Sancier and Fowler (1997), where phonetic drift in the speech of a Portuguese-English bilingual was indicative of a cross-language linkage between the short-lag voiceless stops of Portuguese and the long-lag voiceless stops of English, even though the phonetically closest categories were probably the voiceless stops of Portuguese and the voiced stops of English (both of which are characterized by short-lag VOT). Two studies of L2 vowel production have shown similar resolutions of conflicting phonological and phonetic considerations. In Flege's (1987b) study of French and English speakers, it was found that—in spite of their high degree of phonetic similarity, particularly in the alveolar context that was tested (Strange et al. 2004, 2007; Levy 2009)—French /y/ and English /u/ were not linked to each other for either group. Rather, French /y/ remained unlinked and was produced in a native-like fashion by both groups, while English /u/ was linked to the phonetically more distant, but phonologically parallel French /u/, leading to cross-linguistic interference in the production of these two vowels. This result was replicated by Chang et al. (2010), who demonstrated that the exact same situation obtained with Mandarin and English speakers. When only acoustic measures of vowel quality were considered, English /u/ was shown to be more similar to Mandarin /y/ (which is on the order of 3 Bark away from English /u/ in F_2) than to Mandarin /u/ (which is twice as far away). When the phonological statuses of these vowels were considered, however, Mandarin /u/ emerged as the clear counterpart of English /u/, and production patterns fell in line with the phonological linkage: Mandarin /y/ was produced by L2 learners like a "new" vowel (i.e., close to native phonetic norms), whereas Mandarin /u/ was produced like a "similar" vowel, with interference from English /u/. Taken together, the findings of these studies suggest that when phonetic and phonological considerations conflict with respect to determining cross-linguistic similarity, the higher-level phonological considerations override the lower-level phonetic ones. Categories that are linked at the phonological level in this way are typically similar enough phonetically to be linked at the phonetic level as well, and when this happens the result is assimilatory phonetic drift.

However, it may not always be the case that phonetic proximity is trumped by phonological proximity. For phonological proximity to play a significant role in judgments of cross-linguistic similarity, L2 learners must have sufficient phonological knowledge of the L2. Thus, it may be that novice L2 learners tend not to link L1 and L2 categories in accordance with phonology, since they might not have enough phonological knowledge of the L2 to do so. This hypothesis is supported by the pattern of phonetic drift in English stops discussed in Chapter 4. Recall that cross-linguistic perceptual evidence suggested that of the lenis and fortis Korean stop series, fortis stops were more similar phonetically to English voiced stops in initial position. However, the phonologically closer stop series is arguably the lenis series for two reasons. In Korean, initial lenis stops alternate and are often in free variation with fortis allophones, while the reverse alternation does not occur; in a similar way, initial voiced stops in English are in free variation with voiceless allophones, but not vice versa. Thus, Korean lenis stops and English voiced stops may each be thought of as the "elsewhere" case in processes of laryngeal variation in their respective languages. Korean lenis stops, moreover, are directly associated with phonetically voiced allophones intervocalically. On the basis of these facts, one might expect the English voiced stops to be linked to the Korean lenis stops phonologically. However, the attested pattern, in which English voiced stops did not approximate the long-lag VOT of Korean lenis stops, suggests that, at least at the group level, this particular linkage did not occur. English voiced stops appeared instead to be linked to Korean fortis stops, to which they were so similar in VOT that there was no significant phonetic drift in this dimension.

In short, the present findings indicate that it is inaccurate to claim that phonological proximity always overrides phonetic proximity in determining cross-linguistic similarity. Rather, the interaction of these two metrics is complex and seems to be highly dependent on the extent of learners' phonological knowledge of L2. Given that L2 phonological knowledge deepens with L2 experience, it is claimed that the likelihood of cross-language linkage following phonological information at odds with phonetic information increases with L2 experience. If it can be assumed that cross-language linkages remain malleable over the course of L2 acquisition, this claim leads to the hypothesis that patterns of phonetic drift will change as learners gain more extensive L2 experience and, thus, more phonological knowledge of L2 that might lead to different judgments of cross-linguistic similarity. This hypothesis could be tested in the sort of longitudinal work reported here or in a cross-sectional study examining learners receiving similar input at different points in the L2 acquisition process.

6.3.2 Experience in an Exemplar Model

One of the main empirical contributions of this study is the finding that experience in another language rapidly, and possibly inexorably, alters production of the native language. In Section 6.2.1, this effect was attributed to the restructuring of an L1 exemplar cloud via addition of new L2 exemplars falling in its periphery. In this section, we consider how L2 experience translates to storage of new L2 exemplars.

In order to understand the way in which L2 experience is stored, we must first understand the units in which linguistic experience is stored generally. Commonly the unit of storage is assumed to be the word. In one version of an exemplar model, the episodes of L2 experience stored correspond to word-sized units, and cross-linguistic perceptual differences such as those documented by Best et al. (2001) emerge out of differences in cross-linguistic similarity between lexical items (Johnson 2004). Lexical storage, however, can only be part of the story. The occurrence of L2 influence in L1 production at the level of the phoneme (as seen in Chapter 4) and at the level of the natural class (as seen in Chapter 4, Chapter 5, and Nielsen 2008) implies that there must be a level of abstraction across segment- and class-sized units as well. Thus, English speakers may store experiences of the word *pot* as whole-word episodes, but at some level of organization it must also be possible to access exemplars of /p/ in order for exemplars of L2 words with /p/-like sounds to be able to influence the pronunciation of the /p/ in *pot*. Similarly, exemplars of voiceless plosives must also be accessed together for exemplars of /p/ to be able to influence the production of /k/. Therefore, an exemplar model in which the only units of storage are words must assume that speech production involves activation of a sort that is much more complex than activation of simply the exemplars of the intended word; rather, the process must involve at least activation of exemplars of the intended word, activation of exemplars of words containing phonemes contained in the intended word, and activation of exemplars of words containing phonemes in the natural classes of phonemes contained in the intended word. The alternative is to assume that linguistic experiences are stored in units of multiple sizes, with connections between the various levels. It is beyond the scope of this study to adjudicate between these two approaches; either is consistent with the findings on speech malleability reported in the literature. However, for ease of exposition this study discusses exemplars in terms of exemplars of individual sounds.

Whether or not the units of storage are limited to words, an exemplar model typically includes a way to decide which experiences are ultimately stored. Often this is done via a internal filtering/weighting mechanism to ensure, on the one hand, that aberrant experiences (e.g., speech errors) are discarded, and on the other hand, that canonical experiences (e.g., L1 speech produced in interactive conversation with human beings) are represented more strongly than non-canonical experiences (e.g., artificial speech produced by a computer). One might ask two questions about the workings of this internal mechanism. First, what qualifies as an experience "aberrant" enough not to be stored? Second, what sorts of experiences qualify as non-canonical, and how does the extent of their atypicality affect the strength of their future accessibility?

Normal L2 speech, as fluent human speech, probably does not qualify as an aberrant experience. Most likely, the experience of speech in an L2 currently being learned constitutes canonical experience, except in another language. Aware that they are not hearing L1, L2 learners may store L2 experiences as robust L2 exemplars, which will often fall somewhere within the "cloud" of stored exemplars of the closest L1 sound. The difference between a merged L1-L2 system manifesting cross-linguistic influence and a coexistent L1-L2 system manifesting no cross-linguistic influence might then be conceptualized in terms of a difference in ability to inhibit the activation of exemplars of the irrelevant language. Formulation of a plan for L1 production will lead to activation of only L1 exemplars in a coexistent system, but activation of both L1 exemplars and close L2 exemplars in a merged system, thereby resulting in L2 influence in the L1 output.

The status of ambient L2 exposure is less clear. One might speculate that there is something about *learning* an L2 that causes L1 phonetic drift. Perhaps L2 input needs to be attended to in the way that it is attended to during active learning in order to be stored such that it can affect L1 production. In this case, ambient L2 exposure would be expected to have no effect on L1 production. Some data, however, suggest that over time ambient L2 exposure can have a significant effect on L1 production. In a comparison of VOT in Canadian French and European French, Caramazza and Yeni-Komshian (1974) found that monolingual Canadian French speakers, apparently influenced by the longer VOT of English voiceless stops in the ambient language environment, produced French voiceless stops with longer VOTs than monolingual European French speakers. On the other hand, the English VOTs produced by monolingual Canadian English speakers did not show significant shortening relative to American English VOTs cited in the literature, a result attributed to differences in status between the two languages on both a geopolitical and socioeconomic level. Thirty years later, Fowler, Sramko, Ostry, Rowland, and Hallé (2008) obtained results that were partly consistent with those of Caramazza and Yeni-Komshian (1974), in that they also found no significant difference in VOT between monolingual Canadian English and monolingual American English. However, they found no significant difference in VOT between monolingual Canadian French and monolingual European French, either. This discrepancy with the results of Caramazza and Yeni-Komshian (1974) has been attributed to changes in relative status of the two languages that have transpired in the time intervening between the two studies. Thus, it may be that L2 experience through ambient exposure can in fact be stored in such a way as to result in L1 phonetic drift, but that the effect is mediated by the relative sociolinguistic standing of L1 and L2.

Finally, the status of non-canonical L1 speech, both heard and produced, should also be considered. There is no known literature that has investigated the effects on L1 production of prolonged exposure to non-native speech or of prolonged production of hyperarticulated speech, yet both of these situations obtain in the classroom for language instructors, whose profession involves, at a basic level, trying to make out the communicative intentions behind the often heavily accented speech of their pupils and trying to produce clear and easily comprehensible speech as a model for them (cf. "teacher talk"). The effects of both hearing and pronouncing L1 in this sort of non-canonical way for an extended period of time have not been thoroughly explored. However, given that the non-canonical speech experiences at issue are embedded in the context of meaningful, albeit less-than-fluent, communication, it seems likely that they would be stored somehow, and that an investigation of the speech of language instructors would show significant influence of these experiences in their L1 production.

An analogous effect might occur in the morphosyntactic domain, where it has been observed that the ungrammatical structures most commonly produced by L2 learners can come to sound much less degraded after instructors have been exposed to them numerous times (Iksoo Kwon, p.c.). The reason why this is relevant is that it is widely assumed, by both linguists and by the general public, that language instructors are prototypical examples of speakers of the language that they teach. This assumption is perhaps based on the logic of natural selection: L1 speakers whose speech is far from standard are unlikely to get far as L1 instructors, leaving only L1 speakers with standard, prototypical L1 speech in the profession. It is conceivable that language instructors may initially be selected for in this way. However, future research may show that the accumulation of experience teaching L1 paradoxically causes the L1 production of language instructors to drift away from native norms.

6.3.3 From Individual to Community-Wide Drift

In the study of historical language change, a distinction has been drawn between change at the lexical level (i.e., borrowing of words) and change at a structural level (e.g., phonology, morphology, syntax). Cross-linguistic influence at the lexical level is thought to be relatively common, whereas cross-linguistic influence at a structural level is thought to be less common, requiring a high degree of L2 experience among a large segment of the speech community:

> If there is strong long-term cultural pressure from source-language speakers on the borrowing-language speaker group, then structural features may be borrowed as well [as words]—phonological, phonetic, and syntactic elements, and even (though more rarely) features of the inflectional morphology. Although lexical borrowing frequently takes place without widespread bilingualism, extensive structural borrowing, as has often been pointed out, apparently requires extensive (though **not** universal) bilingualism among borrowing-language speakers over a considerable period of time. (Thomason and Kaufman 1988:37)

Thus, the implication of historical linguistics is that diachronic sound change arising from language contact (Campbell 1976; Boretzky 1991), the kind of change that occurs in a language over generations of speakers, happens typically in conditions of extensive, prolonged contact between L1 and L2. Sapir (1921:193–214) wrote of the progression of such change as "phonetic drift".

In this dissertation, the term "phonetic drift" is used to refer to phonetic change at both a micro level (the idiolect of an individual speaker) and a macro level (the language of a speech community), a terminological conflation that is meant to emphasize the continuity between these two types of change. Simply put, phonetic drift at the micro level may be thought of as planting the seeds of historical sound change. Such contact-induced sound change is likely to occur in areas of high within- and cross-language phonetic variability, in accordance with the claim that "bilingual phonologies may become particularly permeable to inter-linguistic influence precisely where they are acoustically and perceptually unstable, and where they are already congruent to some degree" (Bullock and Gerfen 2004b:103). What is important to note is that the current study suggests, contrary to assumptions made in historical linguistics, that structural change in a domain like phonetics can be significantly accelerated by a relatively brief period of L2 experience.

The propagation of phonetic drift throughout a population might occur in much the same manner as sound changes due to "hypocorrection", one of two complementary mechanisms of change in the theory of sound change proposed by Ohala (1993). Ohala accounted for sound change in terms of non-veridical perception by listeners. Change could result from a perceptual "miss", where a listener failed to correct for features of the acoustic signal resulting from speaker-centric articulatory modifications ("hypocorrection") or from a perceptual "false alarm", where a listener incorrectly interpreted the signal in terms of articulatory modifications that had not in fact occurred ("hypercorrection"). Hypocorrection may provide the link between individual-level phonetic drift and community-level sound change, in that child acquirers of L1 receiving L2-influenced L1 input from a late L2 learner probably do not "correct" for the L2 influence in the L1 speech to which they are exposed. Unless they happen to be acquainted with the talker's linguistic experience (and this plays a role in how the talker's speech is stored), adult interlocutors are also unlikely to normalize for L2 influence in the L1 speech of an ostensibly native L1 talker. Thus, following interactions with L1 talkers manifesting phonetic drift, both child acquirers of L1 and adult monolingual speakers of L1, having failed to correct for the L2 influence, stand to retain L2 influence. In this way, phonetic drift within a few bilingual speakers in one generation may be passed on to later generations of speakers, both bilingual and monolingual.

6.4 Implications for Practice

6.4.1 Teaching Practice

The findings of the current study highlight the continuity of language development across the lifespan. As discussed at length in Chapter 2, L1 production is malleable, and this malleability is reflected in the tendency of L1 to be influenced by experience in L2. This study found phonetic drift in late L2 learners engaged in an intensive program of formal L2 instruction, but whether this result generalizes to more typical foreign language acquisition situations remains to be seen. One might speculate that there was something about the *intensity* of the L2 exposure in the current study that resulted in phonetic drift. Perhaps learners' L1 production drifted because of the four-hour-a-day regimen of L2 instruction, which might have flooded the linguistic system with L2 exemplars that retained relatively strong long-term activation. In this case, one might hypothesize that phonetic drift would not be found in learners acquiring an L2 in a non-intensive elementary foreign language course, which typically consists of an hour of instruction per day at the college level.

Controlled studies of the effects of moderate elementary L2 exposure on L1 production do not exist in the literature, but the SLM predicts that phonetic drift will occur in this case as well, although with a potentially different temporal trajectory. Indeed, it seems rather unlikely that there would be no increase in the long-term activation of L2 exemplars stored over the course of non-intensive language classes. In fact, it is possible that the more distributed nature of the L2 exposure received in a non-intensive course might allow for a greater effect of L2 on L1, as it has been observed by both language learners and language teachers that the content of a language course can often be more effectively learned in a longer non-intensive course than in a shorter intensive course (e.g., Hanna 1969). Thus, there is reason to believe that the present findings would extend to the non-intensive L2 learning situation, but this claim awaits empirical confirmation.

If it turned out to be the case that, on the contrary, phonetic drift only occurred in intensive L2 learning situations, this would be an important finding as well, since it would be indicative of a fundamental difference between the effects of intensive and non-intensive L2 instruction on native language performance. In that possible world, language learners might want to be aware of the fact that intensive instruction, but not non-intensive instruction, is likely to alter their native language production, so that they could weigh the costs and benefits of intensive versus non-intensive classes in a more informed manner. It is not the intention of this study to suggest that maintenance of monolingual-like pronunciation in one's native language should outweigh the advantages of intensive L2 exposure for L2 learning outcomes. Rather, a language-learning public can only benefit from a more accurate understanding of the effects of different kinds of linguistic exposure on linguistic performance, and this understanding should be based on scientific knowledge gained through empirical work on cross-linguistic interaction such as in the current study. Future work on L2 learning should address the gaps in our knowledge of cross-language interaction in order to put together a more complete picture of language development over the lifespan, one that considers multiple sources of linguistic influence on an individual's speech production.

6.4.2 Research Practice

As much as the field of linguistics has accepted the fact that bilinguals are not the sum of two monolinguals, the present findings speak to the need to be ever more careful about defining the subject population for a linguistic study. In studies that concentrate on one language only, it might make sense in some cases to include bilinguals, especially if they are indeed representative of the speech community to which the results are meant to generalize (i.e., a highly diglossic population). However, in most cases it is likely that conflating bilingual participants with monolingual participants obscures, rather than clarifies, our understanding of the object of study.

The empirical problem with using L1 speakers who have some L2 experience as representative of monolingual L1 speakers is particularly evident in studies that make claims about phonetic norms for L1. One example comes from conflicting descriptions of the two-way Swedish voicing contrast between "voiced" and "voiceless" stops by Keating, Linker, and Huffman (1983) and Helgason and Ringen (2008). Keating et al. (1983), who found no voicing during the actual stop closure of Swedish voiced stops, analyzed the Swedish voicing contrast in terms of a short-lag (voiceless unaspirated) category and a long-lag (voiceless aspirated) category, in line with the implementation of this contrast in most other Germanic languages and with a typological preference for economy (i.e., minimal use of phonetic dimensions to establish contrast). The phonetic data of Helgason and Ringen (2008), on the other hand, were indicative of a Swedish voicing system different from that reported by Keating et al. (1983). The Swedish speakers in Helgason and Ringen's study had consistent voicing during the stop closure interval of initial voiced stops, suggesting that the Swedish system is more accurately analyzed as a typologically unusual contrast between (pre)voiced and (post)aspirated stops. Crucially, the Swedish speakers in Helgason and Ringen's study were recorded in Sweden, while those in Keating et al.'s study were recorded in the U.S., after having lived in the U.S. for presumably a considerable period of time. The disparity in results is most likely attributable to this methodological difference between the two studies. As shown in the present study, the Swedish voiced stops of Swedish speakers using English in the U.S. for even a brief period of time could very well have drifted significantly in the direction of English voiced stops (typically implemented as voiceless unaspirated stops in initial position).

It is important to note that the English-language experience of the Swedish speakers in Keating et al.'s (1983) study does not negate the value of their Swedish data. However, the data must be interpreted in accordance with the background characteristics of the participants under study. To be specific, the data of Keating et al. (1983) show something about native Swedish as spoken by Swedish speakers with L2 experience in English, not about native Swedish as spoken by (monolingual) Swedish speakers. The data of Helgason and Ringen (2008), on the other hand, are data that indeed show something about native Swedish as spoken in Sweden.

Thus, the potential for gross disparities in results depending on the L2 experience of participants points to the need for careful control of language background in linguistic studies. The contribution of the present study in this respect is demonstrating that L2 experience must be especially rigorously controlled in studies of speech production, since even brief periods of L2 learning can trigger L1 phonetic drift. For instance, a phonetic study of Vietnamese meant to generalize to monolingual Vietnamese speakers should examine Vietnamese speakers without significant—in particular, recent—L2 experience. In effect, this means that such a study should be conducted in Vietnam, where the ambient language is Vietnamese, instead of in a country like the U.S., where the ambient language is largely English. More generally, this means that linguists who are studying languages that are not native to the surrounding community should expect to work in collaboration with colleagues abroad, since an accurate depiction of these language varieties as they are spoken in their native language environments might not be possible otherwise.

The most rigorous control of linguistic experience, however, is not limited simply to control of foreign language background. Rather, it should extend to control of the *kinds* of linguistic experience individuals have had, both with their native language and with other languages. Experimental studies that are meant to generalize to L1 talkers with typical L1 experience should consider what kinds of linguistic experience are typical. In most countries, for instance, only a small minority of L1 speakers (language instructors) have had extensive experience in a classroom teaching the L1 formally. As discussed in Section 6.3.2, experience teaching an L1 is the sort of experience that could potentially result in modified L1 production (e.g., "teacher talk"), yet it would be difficult to call this experience typical. Hence, teachers of English as a Second Language (ESL), for example, cannot be said to be representative of typical English speakers. As such, ESL teachers constitute one subclass of English speaker that might be reasonably excluded from a study of English, or examined separately and later included in the general group upon confirmation that they pattern with more typical English speakers.

In short, this dissertation should serve as a call to the field for more rigorous control of the linguistic experience of study participants, including fieldwork consultants. Sometimes it is logistically difficult, if not impossible, to find monolingual speakers or to run an experiment in a foreign country. However, the point to take away is not that participants should always be monolingual, but rather that the experiential characteristics of the study sample should accurately represent the population which the study means to investigate.

Chapter 7

Conclusion

7.1 Main Findings

This dissertation documented rapid phonetic drift in L1 as a consequence of elementary experience in an L2. In Chapter 2, it was observed that although evidence from several different research programs suggests that L1 production is highly responsive to the environment, previous studies of cross-linguistic influence assumed that L2 experience could only influence L1 once a high level of L2 proficiency had been reached. The assumption of such a delay in cross-language phonetic interference was shown to be inconsistent with principles of the Speech Learning Model, however, thus motivating the hypothesis that early perceptual linkage between L1 and L2 would cause adult L2 learners to manifest L1 phonetic drift early in L2 acquisition. This hypothesis was supported by the results of the acoustic case studies discussed in Chapters 4 and 5, which showed significant changes in English production as early as Week 2 in the Korean language program.

Chapter 2 also articulated two corollary hypotheses, the first predicting that adult L2 learners would manifest L1 phonetic drift that was assimilatory to L2. This hypothesis was supported by evidence in Chapters 4 and 5 as well. In Chapter 4, it was found that English voiceless stops increased in VOT, approaching the longer VOT of the similar Korean aspirated stops over time. In addition, onset f_0 in English was observed to increase in approximation to the higher onset f_0 of Korean fortis and aspirated stops. Instead of being limited to the corresponding English voiced and voiceless stops, however, the increase in English f_0 occurred more generally, affecting both stop-initial and vowel-initial words. The nature of cross-linguistic assimilation in Chapter 5 was similarly general. In this case, changes in the production of English vowels could not be explained straightforwardly in terms of assimilation to individual Korean vowels; rather, the English vowel space as a whole was found to shift in the F_1 dimension in approximation to the mean F_1 of the model Korean vowel space.

The second corollary hypothesis was that L1 phonetic drift would occur at a level generalizing across segments. Thus, in Chapter 4 it was expected that VOT drift would occur similarly across stops of the same laryngeal specification, regardless of place of articulation. The results of the consonant study were consistent with this hypothesis, but showed that cross-language linkages at the level of the segment also affected the amount of phonetic drift. On the one hand, VOT drift occurred in English voiceless stops at every place of articulation including the alveolars, which differed from Korean aspirated alveolar stops by less than the JND for VOT reported in the literature. This result thus suggests that there was generalization of VOT drift. On the other hand, the amount of VOT drift found differed across places of articulation. Less drift was found in alveolar stops than in bilabial or velar stops, in accordance with differences in VOT norms between English and Korean stops at each place of articulation. Therefore, while there was indeed generalization, phonetic drift remained sensitive to differences between L1 and L2 phonetic norms for parallel categories at the segmental level and not simply at the laryngeal level (i.e., the natural class of segments with the same laryngeal specification). The results of the vowel study reported in Chapter 5 also provided support for the hypothesis of generalized drift, as phonetic drift in vowel formants was found to occur uniformly over the vowel system, rather than diversely for different individual vowels.

Thus, for the L2 learners examined in this study, L2 learning was found to promptly affect L1 production, in a manner that both generalized across phonemic categories and approximated the phonetic properties of L2. Significantly, these results obtained in spite of the study's focus on adult learners, formal speech, and acquisition of an L2 with relatively little in common with the L1—all factors that were thought to reduce the likelihood of finding L2 influence on L1 at an early point in L2 acquisition. The fact that drift occurred even in these ostensibly adverse conditions suggests that there is nothing out of the ordinary about phonetic drift, and that thinking of this phenomenon in terms of attrition misses the bigger picture. Rather than being symptomatic of attrition, phonetic drift seems to be indicative of a fluid, multifaceted quality to language development over the lifespan, wherein production of a particular language can be "nurtured" not only by experience in that language, but by experience in all languages within an individual's linguistic system (which is constantly changing over time). Links between experiences in unrelated languages follow from the inevitable similarities that will exist between languages varying within certain limits, limits that may be imposed by "nature".

The present findings, while consistent with principles of the SLM and the PAM-L2, are only partly predicted by these models, which limit themselves to the scope of cross-linguistic influence at the segmental level. A complete model of cross-linguistic phonetic influence must also provide for the possibility of cross-language developments at a non-segmental level, which, as shown in the current study, constitute a large part of cross-language phonetic effects. Such a model will need to acknowledge the multiple sources of information and influence in speech production, including general mechanisms of speech motor control and somatosensory feedback.

7.2 Future Directions

Though the basic finding of this dissertation regarding L2 experience is clear, the study has raised more questions that it has answered. Chapter 6 identified several areas that are still poorly understood. Future work regarding the nature of phonetic drift should investigate the effect of directional asymmetries in perception on patterns of drift, the effect of somatosensory feedback on the magnitude of drift in different kinds of phonetic dimensions, the extent of cross-linguistic dissimilarity required for drift not to occur, and the occurrence of cross-language dissimilation on a non-segmental level. Future work on the time course of phonetic drift should aim to determine when shifts in L1 production level off, and whether L2-influenced L1 production can drift back to monolingual-like values (and if so, under what circumstances). In this respect, studies of the effects on L1 of distant L2 experience gained in adulthood are very much in order. Individuals with this sort of linguistic experience, often excluded from research because they qualify as neither monolingual nor bilingual, should be studied in their own right—much as heritage speakers are starting to be—in order to better our understanding of linguistic experience: how it is acquired, stored, and forgotten (or not forgotten).

This study has also suggested a way to improve our understanding of the nature of cross-linguistic similarity—in particular, of the interaction between phonetic proximity and phonological proximity. As mentioned in Chapter 6, research on the effects of increasing L2 phonological knowledge on patterns of phonetic drift would help to clarify whether there is a natural hierarchy between phonetic and phonological considerations in the determination of cross-linguistic similarity. It was hypothesized that deeper phonological knowledge of the L2 would result in phonology overriding phonetics in cases of conflict. This hypothesis could be tested, for example, in an artificial language learning experiment in which participants' "phonological" knowledge of another language (e.g., correspondences between sounds in L1 and the artificial language) is manipulated using methods such as explicit instruction.

The theme of this dissertation has been the role of language experience in speech production, and Best and Tyler (2007:18) sum up nicely the kinds of investigations that are necessary for a broader understanding of experiential effects:

> Comparisons among L2, native, unfamiliar native (dialects) and nonnative speech perception would help elucidate how phonological and phonetic information is organized within, between, and beyond the listeners' language(s), and would help determine how that organization changes with increasing knowledge and use of another language or dialect, thus elucidating the effects of language experience on speech perception.

The only thing we should add is that in order to fully understand effects of experience, future research should investigate not only cross-linguistic perception, but also cross-linguistic production. Such work has the potential to contribute to a holistic picture of how L1 speech production can be affected by a wide range of linguistic experiences, including ambient L2 exposure, consistently non-canonical L1 production, interactions with non-native speakers, and L2 learning in the context of non-intensive foreign language classes.

In closing, it should be reiterated that the sorts of L1 phonetic developments documented in this dissertation appear to be entirely normal and are probably much more common than is reflected in the literature on L2 speech. The L2 learners under study did not lose L1 fluency to any detectable degree while undergoing phonetic drift, nor was their casual speech in regular conversations noticeably accented. For this reason, use of the term "attrition" to describe the phenomenon of L1 phonetic drift has been deliberately avoided, as it is a misnomer. Individuals undergoing L1 attrition experience a deterioration in their L1 production as communication is accomplished increasingly in another language, while individuals undergoing phonetic drift experience a change, but not necessarily a deterioration, in their L1 production due to the accumulation of experience in another language. Thus, if there is one thing that can be taken away from the current findings, it is that L1 phonetic drift during L2 acquisition happens as a matter of course. It is much closer to the rule than the exception.

References

Abberton, E. (1972). Some laryngographic data for Korean. *Journal of the International Phonetic Association 2*(2), 67–78.

Ahn, H. (1999). *Post-Release Phonatory Processes in English and Korean: Acoustic Correlates and Implications for Korean Phonology.* Ph. D. thesis, University of Texas at Austin, Austin, TX.

Albert, M. L. and L. K. Obler (1978). *The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism.* New York, NY: Academic Press.

Andersen, R. W. (1982). Determining the linguistic attributes of language attrition. In R. D. Lambert and B. F. Freed (Eds.), *The Loss of Language Skills*, pp. 83–118. Rowley, MA: Newbury House.

Andrews, D. R. (1999). *Sociocultural Perspectives on Language Change in Diaspora: Soviet Immigrants in the United States.* Amsterdam, The Netherlands: John Benjamins Publishing.

Aoyama, K., J. E. Flege, S. G. Guion, R. Akahane-Yamada, and T. Yamada (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics 32*(2), 233–250.

Au, T. K., L. M. Knightly, S.-A. Jun, and J. S. Oh (2002). Overhearing a language during childhood. *Psychological Science 13*(3), 238–243.

Au, T. K. and J. S. Oh (2009). Korean as a heritage language. In C. Lee, G. B. Simpson, and Y. Kim (Eds.), *Handbook of East Asian Psycholinguistics, Volume III: Korean*, pp. 268–275. Cambridge, UK: Cambridge University Press.

Au, T. K., J. S. Oh, L. M. Knightly, S.-A. Jun, and L. F. Romo (2008). Salvaging a childhood language. *Journal of Memory and Language 58*(4), 998–1011.

Au, T. K. and L. F. Romo (1997). Does childhood language experience help adult learners? In H.-C. Chen (Ed.), *The Cognitive Processing of Chinese and Related Asian Languages*, pp. 417–443. Hong Kong: Chinese University Press.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R.* Cambridge, UK: Cambridge University Press.

Babel, M. (2009a). The phonetic and phonological effects of obsolescence in Northern Paiute. In D. Preston and J. Stanford (Eds.), *Variation in Indigenous Minority Languages*, pp. 23–45. Amsterdam, The Netherlands: John Benjamins Publishing.

Babel, M. E. (2009b). *Phonetic and Social Selectivity in Speech Accommodation.* Ph. D. thesis, University of California, Berkeley, Berkeley, CA.

Baker, C. and S. P. Jones (1998). *Encyclopedia of Bilingualism and Bilingual Education.* Clevedon, UK: Multilingual Matters.

Baker, W. and P. Trofimovich (2005). Interaction of native and second language vowel system(s) in early and late bilinguals. *Language and Speech 48*(1), 1–27.

Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, pp. 289–304. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman and H. C. Nusbaum (Eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, pp. 167–224. Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 171–204. Baltimore, MD: York Press.

Best, C. T. and G. W. McRoberts (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech 46*(2–3), 183–216.

Best, C. T., G. W. McRoberts, and E. Goodell (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America 109*(2), 775–794.

Best, C. T., G. W. McRoberts, R. LaFleur, and J. Silver-Isenstadt (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development 18*(3), 339–350.

Best, C. T., G. W. McRoberts, and N. M. Sithole (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by

English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance 14*(3), 345–360.

Best, C. T., A. Traill, A. Carter, K. D. Harrison, and A. Faber (2003). !Xóõ click perception by English, Isizulu, and Sesotho listeners. In M.-J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, pp. 853–856. Causal Productions.

Best, C. T. and M. D. Tyler (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn and M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, pp. 13–34. Amsterdam, The Netherlands: John Benjamins Publishing.

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences 7*(2), 173–221.

Boersma, P. and D. Weenink (2008). Praat: Doing phonetics by computer. Version 5.0.26. http://www.praat.org.

Bohn, O.-S. and J. E. Flege (1996). Perception and production of a new vowel category by adult second language learners. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 53–73. Berlin, Germany: Mouton de Gruyter.

Boretzky, N. (1991). Contact-induced sound change. *Diachronica 8*(1), 1–15.

Bourhis, R. Y. and H. Giles (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, Ethnicity and Intergroup Relations*, pp. 119–136. London, UK: Academic Press.

Bradlow, A. R., R. E. Baker, A. Choi, M. Kim, and K. J. Van Engen (2007). The Wildcat corpus of native- and foreign-accented English. *Journal of the Acoustical Society of America 121*(5), 3072.

Brière, E. J. (1966). An investigation of phonological interference. *Language 42*(4), 768–796.

Broselow, E., S.-I. Chen, and C. Wang (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition 20*(2), 261–280.

Bullock, B. E., A. Dalola, and C. Gerfen (2006). Mapping the patterns of maintenance versus merger in bilingual phonology: The preservation of [a] vs. [ɑ] in Frenchville French. In J.-P. Y. Montreuil (Ed.), *New Perspectives on Romance Linguistics*, Volume II: Phonetics, Phonology and Dialectology, pp. 15–30. Amsterdam, The Netherlands: John Benjamins Publishing.

Bullock, B. E. and C. Gerfen (2004a). Frenchville French: A case study in phonological attrition. *International Journal of Bilingualism 8*(3), 303–320.

Bullock, B. E. and C. Gerfen (2004b). Phonological convergence in a contracting language variety. *Bilingualism: Language and Cognition 7*(2), 95–104.

Bullock, B. E. and C. Gerfen (2005). The preservation of schwa in the converging phonological system of Frenchville French. *Bilingualism: Language and Cognition 8*(2), 117–130.

Bullock, B. E., A. J. Toribio, K. A. Davis, and C. G. Botero (2004). Phonetic convergence in bilingual Puerto Rican Spanish. In V. Chand, A. Kelleher, A. J. Rodríguez, and B. Schmeiser (Eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, Somerville, MA, pp. 113–125. Cascadilla Press.

Burnett, T. A., M. B. Freedland, C. R. Larson, and T. C. Hain (1998). Voice $F0$ responses to manipulations in pitch feedback. *Journal of the Acoustical Society of America 103*(6), 3153–3161.

Burnett, T. A., J. E. Senner, and C. R. Larson (1997). Voice F_0 responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice 11*(2), 202–211.

Campbell, L. (1976). Language contact and sound change. In W. M. Christie, Jr. (Ed.), *Current Progress in Historical Linguistics*, pp. 181–194. Amsterdam, The Netherlands: North-Holland.

Campbell, L. and M. C. Muntzel (1989). The structural consequences of language death. In N. C. Dorian (Ed.), *Investigating Obsolescence: Studies in Language Contraction and Death*, pp. 181–196. Cambridge, UK: Cambridge University Press.

Campbell, R. N. and J. W. Rosenthal (2000). Heritage languages. In J. W. Rosenthal (Ed.), *Handbook of Undergraduate Second Language Education*, pp. 165–184. Mahwah, NJ: Lawrence Erlbaum Associates.

Caramazza, A. and G. H. Yeni-Komshian (1974). Voice onset time in two French dialects. *Journal of Phonetics 2*(3), 239–245.

Caramazza, A., G. H. Yeni-Komshian, E. B. Zurif, and E. Carbone (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America 54*(2), 421–428.

Catford, J. C. and D. B. Pisoni (1970). Auditory vs. articulatory training in exotic sounds. *The Modern Language Journal 54*(7), 477–481.

Chang, C. B. (2009a). Convergence and divergence in language obsolescence. In M. Pak (Ed.), *Current Issues in Unity and Diversity of Languages*, pp. 933–952. Seoul, South Korea: Linguistic Society of Korea.

Chang, C. B. (2009b). English loanword adaptation in Burmese. *Journal of the Southeast Asian Linguistics Society 1*, 77–94.

Chang, C. B. (2009c). Perception vs. production in the development of L2 phonological categories. Paper presented at the 2009 Second Language Research Forum, Michigan State University, East Lansing, MI.

Chang, C. B. (in press). Phonetics vs. phonology in loanword adaptation: Revisiting the role of the bilingual. In A. Bratkievich, D. Bruhn, A. M. Campbell, R. Escamilla, L. Newbold, and R. Rhodes (Eds.), *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Information Structure*, Berkeley, CA. Berkeley Linguistics Society.

Chang, C. B., E. F. Haynes, Y. Yao, and R. Rhodes (2009). A tale of five fricatives: Consonantal contrast in heritage speakers of Mandarin. *University of Pennsylvania Working Papers in Linguistics 15*(1), 37–43.

Chang, C. B., E. F. Haynes, Y. Yao, and R. Rhodes (2010). The phonetic space of phonological categories in heritage speakers of Mandarin. In M. Bane, J. Bueno, T. Grano, A. Grotberg, and Y. McNabb (Eds.), *Proceedings from the 44th Annual Meeting of the Chicago Linguistic Society: The Main Session*, Chicago, IL, pp. 31–45. Chicago Linguistic Society.

Chartrand, T. L. and J. A. Bargh (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology 76*(6), 893–910.

Chaudron, C. (1983). Simplification of input: Topic reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly 17*(3), 437–458.

Childers, D. G. (1978). *Modern Spectrum Analysis*. New York, NY: John Wiley & Sons.

Cho, T., S.-A. Jun, and P. Ladefoged (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics 30*(2), 193–228.

Cho, T. and P. A. Keating (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics 29*(2), 155–190.

Chomsky, N. (1957). *Syntactic Structures*. The Hague, The Netherlands: Mouton de Gruyter.

Coleman, R. F. and I. W. Markham (1991). Normal variations in habitual pitch. *Journal of Voice 5*(2), 173–177.

Crothers, J. (1978). Typology and universals of vowel systems. In J. H. Greenberg, C. A. Ferguson, and E. A. Moravcsik (Eds.), *Universals of Human Language*, Volume 2: Phonology, pp. 93–152. Stanford, CA: Stanford University Press.

Curtiss, S. (1977). *Genie: A Psycholinguistic Study of a Modern-Day "Wild Child"*. Perspectives in Neurolinguistics and Psycholinguistics. New York, NY: Academic Press.

Dart, S. N. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. *Journal of the Acoustical Society of America 81*(1), 138–147.

de Bot, K. (1992). A bilingual processing model: Levelt's 'Speaking' model adapted. *Applied Linguistics 13*(1), 1–24.

de Groot, A. M. B. (1992). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost and L. Katz (Eds.), *Orthography, Phonology, Morphology, and Meaning*, Chapter 20, pp. 389–412. Amsterdam, The Netherlands: North-Holland.

de Melo-Martín, I. (2004). Firing up the nature/nurture controversy: Bioethics and genetic determinism. *Journal of Medical Ethics 31*(9), 526–530.

Dong, Y., S. Gui, and B. MacWhinney (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition 8*(3), 221–238.

Dupoux, E., Y. Hirose, K. Kakehi, C. Pallier, and J. Mehler (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance 25*(6), 1568–1578.

Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *Journal of the Acoustical Society of America 70*(1), 45–50.

Felton, S. K. (1990). Reverse interference in the speech patterns for initial stops of French-dominant English bilinguals. Master's thesis, Ohio State University, Columbus, OH.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics 17*(1), 1–14.

Ferguson, C. A. (1981). 'Foreigner talk' as the name of a simplified register. *International Journal of the Sociology of Language 18*(28), 9–18.

Fishman, J. A. (1972). Societal bilingualism: Stable and transitional. In A. S. Dil (Ed.), *Language in Sociocultural Change*, Chapter 5, pp. 135–152. Stanford, CA: Stanford University Press.

Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning 30*(1), 117–134.

Flege, J. E. (1987a). A critical period for learning to pronounce foreign languages? *Applied Linguistics 8*(2), 162–177.

Flege, J. E. (1987b). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics 15*(1), 47–65.

Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human Communication and Its Disorders: A Review – 1988*, pp. 224–401. Norwood, NJ: Ablex Publishing.

Flege, J. E. (1991a). Age of learning affects the authenticity of voice onset time (VOT) of stop consonants produced in a second language. *Journal of the Acoustical Society of America 89*(1), 395–411.

Flege, J. E. (1991b). The interlingual identification of Spanish and English vowels: Orthographic evidence. *Quarterly Journal of Experimental Psychology 43A*(3), 701–731.

Flege, J. E. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In R. D. Kent (Ed.), *Intelligibility in Speech Disorders: Theory, Measurement, and Management*, Studies in Speech Pathology and Clinical Linguistics, Chapter 5, pp. 157–232. Amsterdam, The Netherlands: John Benjamins Publishing.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 233–272. Baltimore, MD: York Press.

Flege, J. E. (1996). English vowel productions by Dutch talkers: More evidence for the "similar" vs "new" distinction. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 11–52. Berlin, Germany: Mouton de Gruyter.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, and A. Rohde (Eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*, pp. 217–244. Trier, Germany: Wissenschaftlicher Verlag.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole and J. I. Hualde (Eds.), *Laboratory Phonology 9*, pp. 353–382. Berlin, Germany: Walter de Gruyter.

Flege, J. E., O.-S. Bohn, and S. Jang (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics 25*(4), 437–470.

Flege, J. E. and W. Eefting (1987a). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication 6*(3), 185–202.

Flege, J. E. and W. Eefting (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics 15*(1), 67–83.

Flege, J. E. and W. Eefting (1988). Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America 83*(2), 729–740.

Flege, J. E. and J. Hillenbrand (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America 76*(3), 708–721.

Flege, J. E. and I. R. A. MacKay (2010). "Age" effects on second language acquisition. In K. Dziubalska-Kołaczyk, M. Wrembel, and M. Kul (Eds.), *New Sounds 2010: Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech*, Poznań, Poland, pp. 113–118. Adam Mickiewicz University.

Flege, J. E., I. R. A. MacKay, and D. Meador (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America 106*(5), 2973–2987.

Flege, J. E., M. J. Munro, and I. R. A. MacKay (1995a). Effects of age of second-language learning on the production of English consonants. *Speech Communication 16*(1), 1–26.

Flege, J. E., M. J. Munro, and I. R. A. MacKay (1995b). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America 97*(5), 3125–3134.

Flege, J. E., C. Schirru, and I. R. A. MacKay (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication 40*(4), 467–491.

Flege, J. E., G. H. Yeni-Komshian, and S. Liu (1999). Age constraints on second-language acquisition. *Journal of Memory and Language 41*(1), 78–104.

Fokes, J., Z. S. Bond, and M. Steinberg (1985). Acquisition of the English voicing contrast by Arab children. *Language and Speech 28*(1), 81–92.

Forster, J. C. (2008). DMDX. Version 3.2.6.3. http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm.

Fowler, C. A., J. M. Brown, L. Sabadini, and J. Weihing (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language 49*(3), 396–413.

Fowler, C. A., V. Sramko, D. J. Ostry, S. A. Rowland, and P. Hallé (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics 36*(4), 649–663.

Freed, B. F. (1981). Foreigner talk, baby talk, native talk. *International Journal of the Sociology of Language 18*(28), 19–39.

Gass, S. (1984). Development of speech perception and speech production in adult second language learners. *Applied Psycholinguistics 5*(1), 51–74.

Gass, S. M. (1997). *Input, Interaction, and the Second Language Learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M. and E. M. Varonis (1985). Variation in native speaker speech modification to non-native speakers. *Studies in Second Language Acquisition 7*(1), 37–57.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics 15*(2), 87–105.

Giles, H., N. Coupland, and J. Coupland (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, and N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics*, Studies in Emotion and Social Interaction, pp. 1–68. Cambridge, UK: Cambridge University Press.

Godson, L. (2003). *Phonetics of Language Attrition: Vowel Production and Articulatory Setting in the Speech of Western Armenian Heritage Speakers*. Ph.D. dissertation, University of California, San Diego, San Diego, CA.

Godson, L. (2004). Vowel production in the speech of Western Armenian heritage speakers. *Heritage Language Journal 2*(1), 44–69.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review 105*(2), 251–279.

Goldstein, E. B. (2010). *Sensation and Perception* (8th ed.). Belmont, CA: Wadsworth.

Goodfellow, A. and P. Alfred (2002). Maintaining indigenous languages in North America: What can we learn from studies of pidgins and creoles? In B. Burnaby and J. Reyhner (Eds.), *Indigenous Languages Across the Community: Proceedings of the 7th Annual Conference on Stabilizing Indigenous Languages*, Flagstaff, AZ, pp. 212–218. Northern Arizona University.

Goodfellow, A. M. (2005). *Talking in Context: Language and Identity in Kʷakʷa̲ka̲'wakʷ Society*. Montréal, Canada: McGill-Queen's University Press.

Gould, J., H. Lane, J. Vick, J. S. Perkell, M. L. Matthies, and M. Zandipour (2001). Changes in speech intelligibility of postlingually deaf adults after cochlear implantation. *Ear and Hearing 22*(6), 453–460.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition 1*(2), 67–81.

Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior 14*(4), 237–251.

Gregory, S. W. and B. R. Hoyt (1982). Conversation partner mutual adaptation as demonstrated by Fourier series analysis. *Journal of Psycholinguistic Research 11*(1), 35–46.

Grosjean, F. (1982). *Life with Two Languages: An Introduction to Bilingualism*. Cambridge, MA: Harvard University Press.

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One Mind, Two Languages: Bilingual Language Processing*, pp. 1–22. Malden, MA: Blackwell Publishers.

Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica 60*(2), 98–128.

Haggard, M., S. Ambler, and M. Callow (1970). Pitch as a voicing cue. *Journal of the Acoustical Society of America 47*(2B), 613–617.

Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *Journal of the Acoustical Society of America 102*(1), 655–658.

Hallé, P. A. and C. T. Best (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of the Acoustical Society of America 121*(5), 2899–2914.

Hallé, P. A., C. T. Best, and A. Levitt (1999). Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics 27*(3), 281–306.

Hallé, P. A., J. Segui, U. Frauenfelder, and C. Meunier (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance 24*(2), 592–608.

Hammarberg, B. (1996). Conditions on transfer in phonology. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 161–180. Berlin, Germany: Mouton de Gruyter.

Han, J.-I. (1996). *The Phonetics and Phonology of "Tense" and "Plain" Consonants in Korean.* Ph.D. dissertation, Cornell University, Ithaca, NY.

Han, M. S. and R. S. Weitzman (1970). Acoustic features of Korean /P, T, K/, /p, t, k/, and /p^h, t^h, k^h/. *Phonetica 22*(2), 112–128.

Han, N. (1998). A comparative acoustic study of Korean by native Korean children and Korean-American children. Master's thesis, University of California, Los Angeles, Los Angeles, CA.

Hanna, S. A. (1969). Intensive versus non-intensive Arabic. Technical report, University of Utah, Salt Lake City, UT.

Harada, T. (2003). L2 influence on L1 speech in the production of VOT. In M.-J. Solé, D. Recasens, and J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, pp. 1085–1088. Causal Productions.

Hardcastle, W. J. (1973). Some observations on the tense-lax distinction in initial stops in Korean. *Journal of Phonetics 1*(3), 263–272.

Harrison, P. (1996). An experiment with tone. *UCL Working Papers in Linguistics 8*, 575–593.

Hartsuiker, R. J., M. J. Pickering, and E. Veltkamp (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science 15*(6), 409–414.

Hatch, E. M. (1983). *Psycholinguistics: A Second Language Perspective.* Rowley, MA: Newbury House Publishers.

Haynes, E. F. (2010). *Phonetic and Phonological Acquisition in Endangered Languages Learned by Adults: A Case Study of Numu (Oregon Northern Paiute).* Ph. D. thesis, University of California, Berkeley, Berkeley, CA.

Hazan, V., S. Messaoud-Golusi, S. Rosen, S. Nouwens, and B. Shakespeare (2009). Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit? *Journal of Speech, Language, and Hearing Research 52*(6), 1510–1529.

Helgason, P. and C. Ringen (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics 36*(4), 607–628.

Heredia, R. R. and J. M. Brown (2004). Bilingual memory. In T. K. Bhatia and W. C. Ritchie (Eds.), *The Handbook of Bilingualism*, Chapter 9, pp. 225–249. Malden, MA: Blackwell Publishing.

Hillenbrand, J., L. A. Getty, M. J. Clark, and K. Wheeler (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America 97*(5), 3099–3111.

Hinton, L., B. Moonwomon, S. Bremner, H. Luthin, M. van Clay, J. Lerner, and H. Corcoran (1987). It's not just the valley girls: A study of California English. In J. Aske, N. Beery, L. Michaelis, and H. Filip (Eds.), *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Grammar and Cognition*, Berkeley, CA, pp. 117–128. Berkeley Linguistics Society.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review 93*(4), 411–428.

Hirose, H., C. Y. Lee, and T. Ushijima (1974). Laryngeal control in Korean stop production. *Journal of Phonetics 2*(2), 145–152.

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In V. A. Fromkin (Ed.), *Tone: A Linguistic Survey*, pp. 77–111. New York, NY: Academic Press.

Hong, Y. (1987). The e/ä merger in modern Seoul Korean. In S. Kuno et al. (Eds.), *Harvard Studies in Korean Linguistics*, Volume 2, Seoul, South Korea, pp. 367–376. Hanshin Publishing Company.

Houde, J. F. and M. I. Jordan (1998). Sensorimotor adaptation in speech production. *Science 279*(5354), 1213–1216.

Houde, J. F. and M. I. Jordan (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language, and Hearing Research 45*(2), 295–310.

House, A. S. and G. Fairbanks (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America 25*(1), 105–113.

Hudson Kam, C. L. and E. L. Newport (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development 1*(2), 151–195.

Hudson Kam, C. L. and E. L. Newport (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology 59*(1), 30–66.

Huggins, A. W. F. (1972). Just noticeable differences for segment duration in natural speech. *Journal of the Acoustical Society of America 51*(4), 1270–1278.

Hussein, L. H. (1994). *Voicing-Dependent Vowel Duration in Standard Arabic and Its Acquisition by Adult American Students*. Ph.D. dissertation, Ohio State University, Columbus, OH.

Ingram, J. C. L. and S.-G. Park (1996). Inter-language vowel perception and production by Korean and Japanese listeners. In *Proceedings of the 4th International Conference on Spoken Language Processing (INTERSPEECH-1996)*, Philadelphia, PA. International Speech Communication Association.

Ingram, J. C. L. and S.-G. Park (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics 25*(3), 343–370.

International Phonetic Association (1999). *Handbook of the International Phonetic Association*. Cambridge, UK: Cambridge University Press.

Jia, G., W. Strange, Y. Wu, J. Collado, and Q. Guan (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *Journal of the Acoustical Society of America 119*(2), 1118–1130.

Jiang, H. (2008). *Effect of L2 Phonetic Learning on L1 Vowels*. Ph. D. thesis, Simon Fraser University, Vancouver, Canada.

Jiang, H. (2010). Effect of L2 phonetic learning on the production of L1 vowels: A study of Mandarin-English bilinguals in Canada. In K. Dziubalska-Kołaczyk, M. Wrembel, and M. Kul (Eds.), *New Sounds 2010: Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech*, Poznań, Poland, pp. 227–232. Adam Mickiewicz University.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In K. Johnson and J. W. Mullennix (Eds.), *Talker Variability in Speech Processing*, Chapter 8, pp. 145–165. San Diego, CA: Academic Press.

Johnson, K. (2000). Adaptive dispersion in vowel perception. *Phonetica 57*(2–4), 181–188.

Johnson, K. (2004). Cross-linguistic perceptual differences emerge from the lexicon. In A. Agwuele, W. Warren, and S.-H. Park (Eds.), *Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in Speech Production and Perception*, Somerville, MA, pp. 26–41. Cascadilla Proceedings Project.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics 34*(4), 485–499.

Johnson, K. (2008). *Quantitative Methods in Linguistics.* Malden, MA: Blackwell Publishing.

Johnson, K. and M. Babel (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics 38*(1), 127–136.

Jones, J. A. and K. G. Munhall (2000). Perceptual calibration of $F0$ production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America 108*(3), 1246–1251.

Jones, J. A. and K. G. Munhall (2005). Remapping auditory-motor representations in voice production. *Current Biology 15*(19), 1768–1772.

Jones, P. E. (1995). Contradictions and unanswered questions in the Genie case: A fresh look at the linguistic evidence. *Language and Communication 15*(3), 261–280.

Joseph, B. D. (2009). Broad vs. localistic dialectology, standard vs. dialect: The case of the Balkans and the drawing of linguistic boundaries. In S. Tsiplakou, M. Karyolemou, and P. Pavlou (Eds.), *Language Variation – European Perspectives II: Selected Papers from the 4^{th} International Conference on Language Variation in Europe (ICLaVE 4)*, pp. 119–135. Amsterdam, The Netherlands: John Benjamins Publishing.

Jun, S.-A. (1993). *The Phonetics and Phonology of Korean Prosody.* Ph. D. thesis, Ohio State University, Columbus, OH.

Junqua, J.-C. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers. *Journal of the Acoustical Society of America 93*(1), 510–524.

Kagaya, R. (1974). A fiberscopic and acoustic study of the Korean stops, affricates and fricatives. *Journal of Phonetics 2*(2), 161–180.

Kaltenbacher, E. (1997). German speech rhythm in L2 acquisition. In J. Leather and A. James (Eds.), *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second-Language Speech*, Klagenfurt, Austria, pp. 158–166. University of Klagenfurt.

Kang, K.-H. and S. G. Guion (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *Journal of the Acoustical Society of America 119*(3), 1672–1683.

Kang, K.-H. and S. G. Guion (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of the Acoustical Society of America 124*(6), 3909–3917.

Kang, Y. (2008). Interlanguage segmental mapping as evidence for the nature of lexical representation. *Language and Linguistics Compass 2*(1), 103–118.

Katseff, S. and J. Houde (2008). Partial compensation in speech adaptation. *UC Berkeley Phonology Lab Annual Report 2008*, 445–461.

Kawahara, H. (1993). Transformed auditory feedback: Effects of fundamental frequency perturbation. *Journal of the Acoustical Society of America 94*(3), 1883–1884.

Kawahara, H. (1994). Interactions between speech production and perception under auditory feedback perturbations on fundamental frequencies. *Journal of the Acoustical Society of Japan 15*(3), 201–202.

Keating, P., W. Linker, and M. Huffman (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics 11*(3), 277–290.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language 60*(2), 286–319.

Kim, C.-W. (1965). On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). *Word 21*(3), 339–359.

Kim, C.-W. (1968). The vowel system of Korean. *Language 44*(3), 516–527.

Kim, C.-W. (1970). A theory of aspiration. *Phonetica 21*(2), 107–116.

Kim, H.-S. (2008). The phonetics and phonology of vowel length variation in Korean reduplicated ideophones. *Eumseong, Eumun, Hyeongtaeron Yeon'gu [Studies in Phonetics, Phonology and Morphology] 14*(2), 39–60.

Kim, M. (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates. In *Proceedings of the 8^{th} International Conference on Spoken Language Processing (INTERSPEECH-2004)*, Jeju Island, Korea, pp. 49–52. International Speech Communication Association.

Kim, M. (2009). Phonetic accommodation in conversations between native and nonnative speakers. *Journal of the Acoustical Society of America 125*(4), 2764.

Kim, M., A. Bradlow, and W. Horton (2007a). Phonetic convergence and divergence during discourse between native and nonnative speakers. Poster presented at the 2007 Mid-Continental Workshop on Phonology, Ohio State University, Columbus, OH.

Kim, M., A. Bradlow, and W. Horton (2007b). Phonetic convergence between native and nonnative speakers. Paper presented at the 4th Seoul International Conference on Phonology and Morphology, Seoul, South Korea.

Kim, M.-R., P. S. Beddor, and J. Horrocks (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics 30*(1), 77–100.

Kim, M.-R. and S. Duanmu (2004). Tense and lax stops in Korean. *Journal of East Asian Linguistics 13*(1), 59–104.

Kim, M.-R. C. (1994). *Acoustic Characteristics of Korean Stops and Perception of English Stop Consonants*. Ph. D. thesis, University of Wisconsin, Madison, Madison, WI.

Kim, M.-R. C. and A. J. Lotto (2002). An investigation of acoustic characteristics of Korean stops produced by non-heritage learners. In J. J. Ree (Ed.), *The Korean Language in America*, Volume 7, pp. 177–188. American Association of Teachers of Korean.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America 59*(5), 1208–1221.

Knightly, L. M., S.-A. Jun, J. S. Oh, and T. K. Au (2003). Production benefits of childhood overhearing. *Journal of the Acoustical Society of America 114*(1), 465–474.

Ko, I. (2009). The merger of *ey* /e/ and *ay* /ɛ/ of Seoul Korean. *Eoneo Jeongbo [Language Information] 10*, 73–89.

Kroll, J. F. and E. Stewart (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language 33*(2), 149–174.

Kroll, J. F. and N. Tokowicz (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll and A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*, pp. 531–553. New York, NY: Oxford University Press.

Kuhl, P. K. and P. Iverson (1995). Linguistic experience and the "Perceptual Magnet Effect". In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 121–154. Baltimore, MD: York Press.

Labov, W. (1963). The social motivation of a sound change. *Word 19*(3), 273–309.

Labov, W. (1994). *Principles of Linguistic Change*, Volume 1: Internal Factors. Malden, MA: Blackwell Publishers.

Labov, W., S. Ash, and C. Boberg (2006). *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin, Germany: Mouton de Gruyter.

LaCharité, D. and C. Paradis (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry 36*(2), 223–258.

Ladefoged, P. (1999). American English. In International Phonetic Association (Ed.), *Handbook of the International Phonetic Association*, pp. 41–44. Cambridge, UK: Cambridge University Press.

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor, MI: University of Michigan Press.

Laeufer, C. (1996). Towards a typology of bilingual phonological systems. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 325–342. Berlin, Germany: Mouton de Gruyter.

Lane, H., M. Matthies, J. Perkell, J. Vick, and M. Zandipour (2001). The effects of changes in hearing status in cochlear implant users on the acoustic vowel space and CV coarticulation. *Journal of Speech, Language, and Hearing Research 44*(3), 552–563.

Lane, H., J. Perkell, M. Svirsky, and J. Webster (1991). Changes in speech breathing following cochlear implant in postlingually deafened adults. *Journal of Speech and Hearing Research 34*(3), 526–533.

Lane, H., J. Perkell, J. Wozniak, J. Manzella, P. Guiod, M. Matthies, M. MacCollin, and J. Vick (1998). The effect of changes in hearing status on speech sound level and speech breathing: A study conducted with cochlear implant users and NF-2 patients. *Journal of the Acoustical Society of America 104*(5), 3059–3069.

Lane, H. and B. Tranel (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research 14*(4), 677–709.

Lane, H. and J. W. Webster (1991). Speech deterioration in postlingually deafened adults. *Journal of the Acoustical Society of America 89*(2), 859–866.

Lane, H., J. Wozniak, M. Matthies, M. Svirsky, and J. Perkell (1995). Phonemic resetting versus postural adjustments in the speech of cochlear implant users: An exploration of voice-onset time. *Journal of the Acoustical Society of America 98*(6), 3096–3106.

Lane, H., J. Wozniak, and J. Perkell (1994). Changes in voice-onset time in speakers with cochlear implants. *Journal of the Acoustical Society of America 96*(1), 56–64.

Larson, C. R., K. W. Altman, H. Liu, and T. C. Hain (2008). Interactions between auditory and somatosensory feedback for voice F_0 control. *Experimental Brain Research 187*(4), 613–621.

Larson, C. R., T. A. Burnett, S. Kiran, and T. C. Hain (2000). Effects of pitch-shift velocity on voice F_0 responses. *Journal of the Acoustical Society of America 107*(1), 559–564.

Leather, J. (1996). Interrelation of perceptual and productive learning in the initial acquisition of second-language tone. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 75–101. Berlin, Germany: Mouton de Gruyter.

Lee, B. S. (1950). Some effects of side-tone delay. *Journal of the Acoustical Society of America 22*(5), 639–640.

Lee, H. B. (1993). Korean. *Journal of the International Phonetic Association 23*(1), 28–31.

Lee, J.-H. (1995). The /e/-/æ/ merger in modern Seoul Korean is a 'near' merger. In S. Kuno et al. (Eds.), *Harvard Studies in Korean Linguistics*, Volume 6, Seoul, South Korea, pp. 108–120. Hanshin Publishing Company.

Lee, K.-H. and M.-S. Jung (2000). Acoustic characteristics and perceptual cues for Korean stops. *Eumseong Gwahag [Speech Sciences] 7*(2), 139–155.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lehiste, I. and G. E. Peterson (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America 33*(4), 419–425.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York, NY: John Wiley & Sons.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America 125*(2), 1138–1152.

Lewandowski, N. (2009). Sociolinguistic factors in language proficiency: Phonetic convergence as a signature of pronunciation talent. In G. Dogil and S. M. Reiterer (Eds.), *Language Talent and Brain Activity*, pp. 257–278. Berlin, Germany: Mouton de Gruyter.

Liljencrants, J. and B. Lindblom (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language 48*(4), 839–862.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America 35*(11), 1773–1781.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. J. Ohala and J. J. Jaeger (Eds.), *Experimental Phonology*, pp. 13–44. Orlando, FL: Academic Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle and A. Marchal (Eds.), *Speech Production and Speech Modeling*, pp. 403–439. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lindblom, B. and O. Engstrand (1989). In what sense is speech quantal? *Journal of Phonetics 17*(1–2), 107–121.

Lindblom, B. and I. Maddieson (1988). Phonetic universals in consonant systems. In L. M. Hyman and C. N. Li (Eds.), *Language, Speech, and Mind: Studies in Honour of Victoria A. Fromkin*, pp. 62–78. London, UK: Routledge.

Lisker, L. and A. S. Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word 20*(3), 384–422.

Lisker, L. and A. S. Abramson (1967). Some effects of context on voice onset time in English stops. *Language and Speech 10*(1), 1–28.

Lisker, L., A. M. Liberman, D. M. Erickson, D. Dechovitz, and R. Mandler (1977). On pushing the voice-onset-time (VOT) boundary about. *Language and Speech 20*(3), 209–216.

Locke, J. (1690/1995). *An Essay Concerning Human Understanding*. Amherst, NY: Prometheus Books.

Lombard, É. (1911). Le signe de l'élévation de la voix. *Annales des Maladies de l'Oreille, du Larynx, du Nez et du Pharynx 37*(2), 101–119.

Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition 5*(2), 177–193.

Lorenzen, B. and L. L. Murray (2008). Bilingual aphasia: A theoretical and clinical review. *American Journal of Speech-Language Pathology 17*(3), 299–317.

Luthin, H. W. (1987). The story of California /ow/: The coming-of-age of English in California. In K. M. Dennin, S. Inkelas, F. C. McNair-Knox, and J. R. Rickford (Eds.), *Variation in Language: NWAV-XV at Stanford*, Stanford, CA, pp. 312–324. Stanford Linguistics Association.

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception and Psychophysics 46*(2), 187–200.

Mack, M. (1990). Phonetic transfer in a French-English bilingual child. In P. H. Nelde (Ed.), *Language Attitudes and Language Conflict*, pp. 107–124. Bonn, Germany: Dümmler.

Mack, M. (2003). The phonetic systems of bilinguals. In M. T. Banich and M. Mack (Eds.), *Mind, Brain, and Language: Multidisciplinary Perspectives*, pp. 309–349. Mahwah, NJ: Lawrence Erlbaum Press.

MacKay, I. R. A., J. E. Flege, T. Piske, and C. Schirru (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America 110*(1), 516–528.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge, UK: Cambridge University Press.

Major, R. C. (1987a). A model for interlanguage phonology. In G. Ioup and S. H. Weinberger (Eds.), *Interlanguage Phonology: The Acquisition of a Second Language Sound System*, pp. 101–124. Cambridge, MA: Newbury House Publishers.

Major, R. C. (1987b). Phonological similarity, markedness, and rate of L2 acquisition. *Studies in Second Language Acquisition 9*(1), 63–82.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal 76*(2), 190–208.

Major, R. C. (1996). L2 acquisition, L1 loss, and the critical period hypothesis. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 147–159. Berlin, Germany: Mouton de Gruyter.

Major, R. C. (2001). *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Matthies, M. L., M. Svirsky, J. Perkell, and H. Lane (1996). Acoustic and articulatory measures of sibilant production with and without auditory feedback from a cochlear implant. *Journal of Speech and Hearing Research 39*(5), 936–946.

Matthies, M. L., M. A. Svirsky, H. L. Lane, and J. S. Perkell (1994). A preliminary study of the effects of cochlear implants on the production of sibilants. *Journal of the Acoustical Society of America 96*(3), 1367–1373.

Mayberry, R. I. (1993). First-language acquisition after childhood differs from second-language acquisition: The case of American Sign Language. *Journal of Speech and Hearing Research 36*(6), 1258–1270.

Mayberry, R. I. (2007). When timing is everything: Age of first-language acquisition effects on second-language learning. *Applied Psycholinguistics 28*(3), 537–549.

Mayberry, R. I. (2010). Early language acquisition and adult language ability: What sign language reveals about the critical period for language. In M. Marschark and P. E. Spencer (Eds.), *Oxford Handbook of Deaf Studies, Language, and Education*, Volume 2, pp. 281–291. Oxford, UK: Oxford University Press.

Mayberry, R. I. and E. B. Eichen (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language 30*(4), 486–512.

Mayberry, R. I. and E. Lock (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and Language 87*(3), 369–384.

Mayberry, R. I., E. Lock, and H. Kazmi (2002). Linguistic ability and early language exposure. *Nature 417*(6884), 38.

Mitterer, H. and M. Ernestus (2008). The link between speech perception and production is phonological and abstract: Evidence from a shadowing task. *Cognition 109*(1), 168–173.

Montrul, S. A. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam, The Netherlands: John Benjamins Publishing.

Moore, D. S. (2003). *The Dependent Gene: The Fallacy of "Nature vs. Nurture"*. New York, NY: Henry Holt and Co.

Morford, J. P. and R. I. Mayberry (2000). A reexamination of "early exposure" and its implications for language acquisition by eye. In C. Chamberlain, J. P. Morford, and R. I. Mayberry (Eds.), *Language Acquisition by Eye*, Chapter 7, pp. 111–128. Mahwah, NJ: Lawrence Erlbaum Associates.

Munro, M. J. (1993). Non-native productions of English vowels: Acoustic measurements and accentedness ratings. *Language and Speech 36*(1), 39–66.

Munro, M. J., J. E. Flege, and I. R. A. MacKay (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics 17*(3), 313–334.

Murray, K. R. (1980). Hemispheric and directional asymmetry of pitch discrimination. *Journal of Research in Music Education 28*(4), 225–228.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology 32*(5), 790–804.

Nearey, T. M. and B. L. Rochet (1994). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association 24*(1), 1–18.

Nielsen, K. Y. (2005). Generalization of phonetic imitation across place of articulation. In V. Hazan and P. Iverson (Eds.), *Proceedings of the ISCA Workshop on Plasticity in Speech Perception (PSP2005)*, London, UK, pp. 47–50. International Speech Communication Association.

Nielsen, K. Y. (2006). Specificity and generalizability of spontaneous phonetic imitation. In *Proceedings of the 9^{th} International Conference on Spoken Language Processing (INTERSPEECH-2006)*, Pittsburgh, PA. International Speech Communication Association.

Nielsen, K. Y. (2007a). Implicit phonetic imitation is constrained by phonemic contrast. In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16^{th} International Congress of Phonetic Sciences*, Dudweiler, Germany, pp. 1961–1964. Pirrot.

Nielsen, K. Y. (2007b). The interaction between spontaneous imitation and linguistic knowledge. *UCLA Working Papers in Phonetics 105*, 125–137.

Nielsen, K. Y. (2008). *The Specificity of Allophonic Variability and Its Implications for Accounts of Speech Perception and Production.* Ph. D. thesis, University of California, Los Angeles, Los Angeles, CA.

Oglesbee, E. and D. Kewley-Port (2009). Estimating vowel formant discrimination thresholds using a single-interval classification task. *Journal of the Acoustical Society of America 125*(4), 2323–2335.

Oh, J., S.-A. Jun, L. Knightly, and T. Au (2003). Holding on to childhood language memory. *Cognition 86*(3), B53–B64.

Oh, J. S. and T. K. Au (2005). Learning Spanish as a heritage language: The role of sociocultural background variables. *Language, Culture and Curriculum 18*(3), 229–241.

Oh, J. S., T. K. Au, and S.-A. Jun (2002). Benefits of childhood language experience for adult L2 learners' phonology. In B. Skarabela, S. Fish, and A. H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, Volume 2, Somerville, MA, pp. 464–472. Cascadilla Press.

Oh, M. and K. Johnson (1997). A phonetic study on Korean intervocalic laryngeal consonants. *Eumseong Gwahag [Speech Sciences] 1*(1), 83–101.

Ohala, J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical Linguistics: Problems and Perspectives*, Chapter 9, pp. 237–278. London, UK: Addison Wesley Publishing.

Pallier, C., S. Dehaene, J.-B. Poline, D. LeBihan, A.-M. Argenti, E. Dupoux, and J. Mehler (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex 13*(2), 155–161.

Paradis, C. and D. LaCharité (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics 33*(2), 379–430.

Paradis, M. (2001). An integrated neurolinguistic theory of bilingualism (1976-2000). *LACUS Forum 27*, 5–15.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America 119*(4), 2382–2393.

Pardo, J. S. (2009). Expressing oneself in conversational interaction. In E. Morsella (Ed.), *Expressing Oneself / Expressing One's Self: Communication, Cognition, Language, and Identity.* London, UK: Psychology Press.

Park, J.-W. (1994). Variation of vowel length in Korean. In Y.-K. Kim-Renaud (Ed.), *Theoretical Issues in Korean Linguistics*, pp. 175–187. Stanford, CA: Center for the Study of Language and Information.

Patel, R. and K. W. Schell (2008). The influence of linguistic content on the Lombard effect. *Journal of Speech, Language, and Hearing Research 51*(1), 209–220.

Penfield, W. and L. Roberts (1959). *Speech and Brain Mechanisms.* New York, NY: Atheneum.

Peng, S. (1993). Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. *Phonetica 50*(4), 245–260.

Perkell, J., H. Lane, M. Svirsky, and J. Webster (1992). Speech of cochlear implant patients: A longitudinal study of vowel production. *Journal of the Acoustical Society of America 91*(5), 2961–2978.

Perkell, J., M. Matthies, H. Lane, F. Guenther, R. Wilhelms-Tricarico, J. Wozniak, and P. Guiod (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication 22*(2–3), 227–250.

Perkell, J., W. Numa, J. Vick, H. Lane, T. Balkany, and J. Gould (2001). Language-specific, hearing-related changes in vowel spaces: A preliminary study of English- and Spanish-speaking cochlear implant users. *Ear and Hearing 22*(6), 461–470.

Peterson, G. E. and H. L. Barney (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America 24*(1), 175–184.

Pickering, M. J. and S. Garrod (2004a). The interactive alignment model: Developments and refinements. *Behavioral and Brain Sciences 27*(2), 212–219.

Pickering, M. J. and S. Garrod (2004b). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences 27*(2), 169–190.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (Eds.), *Frequency Effects and the Emergence of Lexical Structure*, pp. 137–157. Amsterdam, The Netherlands: John Benjamins Publishing.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven and N. Warner (Eds.), *Laboratory Phonology 7*, pp. 101–139. Berlin, Germany: Mouton de Gruyter.

Pile, E. J. S., H. R. Dajani, D. W. Purcell, and K. G. Munhall (2007). Talking under conditions of altered auditory feedback: Does adaptation of one vowel generalize to other vowels? In J. Trouvain and W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, Dudweiler, Germany, pp. 645–648. Pirrot.

Pinheiro, J. C. and D. M. Bates (2000). *Mixed-Effects Models in S and S-PLUS*. New York, NY: Springer Verlag.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal 6*(1), 40–71.

Polka, L. and O.-S. Bohn (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America 100*(1), 577–592.

Port, R. F. and F. M. Mitleb (1983). Segmental features and implementation in acquisition of English by Arabic speakers. *Journal of Phonetics 11*(3), 219–229.

Potter, M. C., K.-F. So, B. Von Eckardt, and L. B. Feldman (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior 23*(1), 23–38.

Purcell, D. W. and K. G. Munhall (2006). Compensation following real-time manipulation of formants in isolated vowels. *Journal of the Acoustical Society of America 119*(4), 2288–2297.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramscar, M. and D. Yarlett (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science 31*(6), 927–960.

Ridley, M. (2003). *Nature via Nurture: Genes, Experience, and What Makes Us Human*. New York, NY: HarperCollins Publishers.

Roche, J. (1998). Variation in xenolects (foreigner talk). In U. Ammon, K. J. Mattheier, and P. H. Nelde (Eds.), *Sociolinguistica*, Volume 12, pp. 117–139. Tübingen, Germany: Max Niemeyer Verlag.

Roederer, J. G. (1973). *Introduction to the Physics and Psychophysics of Music*. London, UK: English Universities Press.

Saffran, J. R., R. N. Aslin, and E. L. Newport (1996). Statistical learning by 8-month-old infants. *Science 274*(5294), 1926–1928.

Saffran, J. R., E. L. Newport, and R. N. Aslin (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language 35*(4), 606–621.

Sancier, M. L. and C. A. Fowler (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics 27*(4), 421–436.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York, NY: Harcourt, Brace and Co.

Schmidt, A. M. (2007). Cross-language consonant identification: English and Korean. In O.-S. Bohn and M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, Chapter 11, pp. 185–200. Amsterdam, The Netherlands: John Benjamins Publishing.

Schwartz, J.-L., L.-J. Boë, N. Vallée, and C. Abry (1997). The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics 25*(3), 255–286.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning 19*(3), 245–253.

Sebastián-Gallés, N. and S. Soto-Faraco (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition 72*(2), 111–123.

Seidenberg, M. S. and J. L. McClelland (1989). A distributed developmental model of word recognition and naming. *Psychological Review 96*(4), 523–568.

Seliger, H. W. and R. M. Vago (1991). The study of first language attrition: An overview. In H. W. Seliger and R. M. Vago (Eds.), *First Language Attrition*, pp. 3–15. Cambridge, UK: Cambridge University Press.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching 10*(3), 209–231.

Selishchev, A. M. (1925). Des traits linguistiques communes aux langues balkaniques: Un balkanisme ancien en bulgare. *Revue des Études Slaves 5*, 38–57.

Sheldon, A. and W. Strange (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics 3*(3), 243–261.

Shepard, C. A., H. Giles, and B. A. LePoire (2001). Communication accommodation theory. In W. P. Robinson and H. Giles (Eds.), *The New Handbook of Language and Social Psychology*, pp. 33–56. Chichester, UK: John Wiley & Sons.

Shin, E. (2007). How do non-heritage students learn to make the three-way contrast of Korean stops? In H.-S. Wang (Ed.), *The Korean Language in America*, Volume 12, pp. 85–105. American Association of Teachers of Korean.

Shockley, K., L. Sabadini, and C. A. Fowler (2004). Imitation in shadowing words. *Perception and Psychophysics 66*(3), 422–429.

Silva, D. J. (1992). *The Phonetics and Phonology of Stop Lenition in Korean.* Ph. D. thesis, Cornell University, Ithaca, NY.

Silva, D. J. (2006a). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology 23*(2), 287–308.

Silva, D. J. (2006b). Variation in voice onset time for Korean stops: A case for recent sound change. *Korean Linguistics 13*, 1–16.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass 3*(2), 621–640.

Sohn, H.-M. (1999). *The Korean Language.* Cambridge Language Surveys. Cambridge, UK: Cambridge University Press.

Spilka, B. (1954). Relationships between certain aspects of personality and some vocal effects of delayed speech feedback. *Journal of Speech and Hearing Disorders 19*(4), 491–503.

Strange, W. (1992). Learning non-native phoneme contrasts: Interactions among stimulus, subject, and task variables. In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Eds.), *Speech Perception, Production and Linguistic Structure*, pp. 197–219. Tokyo, Japan: Ohmsha.

Strange, W., E. Levy, and R. Lehnholf, Jr. (2004). Perceptual assimilation of French and German vowels by American English monolinguals: Acoustic similarity does not predict perceptual similarity. *Journal of the Acoustical Society of America 115*(5), 2606.

Strange, W., A. Weber, E. S. Levy, V. Shafiro, M. Hisagi, and K. Nishi (2007). Acoustic variability within and across German, French and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America 122*(2), 1111–1129.

Summers, W. V., D. B. Pisoni, R. H. Bernacki, R. I. Pedlow, and M. A. Stokes (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America 84*(3), 917–928.

Svirsky, M. A., H. Lane, J. S. Perkell, and J. Wozniak (1992). Effects of short-term auditory deprivation on speech production in adult cochlear implant users. *Journal of the Acoustical Society of America 92*(3), 1284–1300.

Tahta, S., M. Wood, and K. Loewenthal (1981). Age changes in the ability to replicate foreign language pronunciation and intonation. *Language and Speech 24*(4), 363–372.

Tees, R. C. and J. F. Werker (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology 38*(4), 579–590.

Thomason, S. G. (2001). *Language Contact: An Introduction.* Washington, DC: Georgetown University Press.

Thomason, S. G. and T. Kaufman (1988). *Language Contact, Creolization, and Genetic Linguistics.* Berkeley, CA: University of California Press.

Toribio, A. J., B. E. Bullock, C. G. Botero, and K. A. Davis (2005). Perseverative phonetic effects in bilingual code-switching. In R. S. Gess and E. J. Rubin (Eds.), *Theoretical and Experimental Approaches to Romance Linguistics*, pp. 291–306. Amsterdam, The Netherlands: John Benjamins Publishing.

Tremblay, S., D. M. Shiller, and D. J. Ostry (2003). Somatosensory basis of speech production. *Nature 423*(6942), 866–869.

Trudgill, P. and T. Foxcroft (1978). On the sociolinguistics of vocalic mergers: Transfer and approximation in East Anglia. In P. Trudgill (Ed.), *Sociolinguistic Patterns in British English*, pp. 69–79. London, UK: Edward Arnold.

van Hell, J. G. and A. M. B. de Groot (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition 1*(3), 193–211.

Ventureyra, V. A. G. and C. Pallier (2004). In search of the lost language: The case of adopted Koreans in France. In M. S. Schmid, B. Köpke, M. Keijzer, and L. Weilemar (Eds.), *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*, pp. 207–221. Amsterdam, The Netherlands: John Benjamins Publishing.

Ventureyra, V. A. G., C. Pallier, and H.-Y. Yoo (2004). The loss of first language phonetic perception in adopted Koreans. *Journal of Neurolinguistics 17*(1), 79–91.

Vick, J. C., H. Lane, J. S. Perkell, M. L. Matthies, J. Gould, and M. Zandipour (2001). Covariation of cochlear implant users' perception and production of vowel contrasts and their identification by listeners with normal hearing. *Journal of Speech, Language, and Hearing Research 44*(6), 1257–1267.

von Békésy, G. (1960). *Experiments in Hearing*. New York, NY: McGraw-Hill.

Ward, N., M. Sundara, B. Conboy, and P. Kuhl (2009). Consequences of short-term language exposure in infancy on babbling. *Journal of the Acoustical Society of America 126*(4), 2311.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. The Hague, The Netherlands: Mouton de Gruyter.

Weinreich, U. (1957). On the description of phonic interference. *Word 13*(1), 1–11.

White, L. (1989). *Universal Grammar and Second Language Acquisition*. Language Acquisition and Language Disorders. Amsterdam, The Netherlands: John Benjamins Publishing.

Whiteside, S. P. (1998a). Identification of a speaker's sex: A study of vowels. *Perceptual and Motor Skills 86*(2), 579–584.

Whiteside, S. P. (1998b). The identification of a speaker's sex from synthesized vowels. *Perceptual and Motor Skills 87*(2), 595–600.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception and Psychophysics 21*(4), 289–297.

Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception and Psychophysics 26*(2), 95–104.

Winkworth, A. L. and P. J. Davis (1997). Speech breathing and the Lombard effect. *Journal of Speech, Language, and Hearing Research 40*(1), 159–169.

Wode, H. (1978). The beginning of non-school room L2 phonological acquisition. *International Review of Applied Linguistics in Language Teaching 16*(2), 109–125.

Wode, H. (1994). Nature, nurture, and age in language acquisition: The case of speech perception. *Studies in Second Language Acquisition 16*(3), 325–345.

Woolard, K. A. (1989). Language convergence and language death as social processes. In N. C. Dorian (Ed.), *Investigating Obsolescence: Studies in Language Contraction and Death*, pp. 355–368. Cambridge, UK: Cambridge University Press.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds: Perception of American English /r/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 305–320. Baltimore, MD: York Press.

Yamada, R. A. and Y. Tohkura (1992). Perception of American English /r/ and /l/ by native speakers of Japanese. In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka (Eds.), *Speech Perception, Production and Linguistic Structure*, pp. 155–174. Tokyo, Japan: Ohmsha.

Yamada, R. A., Y. Tohkura, and N. Kobayashi (1996). Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/, and /w/ by native speakers of Japanese. In A. James and J. Leather (Eds.), *Second-Language Speech: Structure and Process*, pp. 103–117. Berlin, Germany: Mouton de Gruyter.

Yang, B. (1992). An acoustical study of Korean monophthongs produced by male and female speakers. *Journal of the Acoustical Society of America 91*(4), 2280–2283.

Yang, B. (1996a). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics 24*(2), 245–261.

Yang, B. (1996b). Perceptual contrast in the Korean and English vowel system normalized. In *Proceedings of the 4^{th} International Conference on Spoken Language Processing (INTERSPEECH-1996)*, Volume 1, Philadelphia, PA. International Speech Communication Association.

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin 60*(3), 213–232.

Yeni-Komshian, G. H., J. E. Flege, and S. Liu (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition 3*(2), 131–149.

Yusa, N., K. Nasukawa, M. Koizumi, J. Kim, N. Kimura, and K. Emura (2010). Unexpected effects of the second language on the first. In K. Dziubalska-Kołaczyk, M. Wrembel, and M. Kul (Eds.), *New Sounds 2010: Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech*, Poznań, Poland, pp. 580–584. Adam Mickiewicz University.

Appendix A

Entrance Questionnaire

Instructions: *The purpose of this questionnaire is to gather some information about you and your language background. Please fill it out as completely and accurately as possible.* ***All information will be kept strictly anonymous.*** *You may not remember enough about your early childhood to answer certain questions; in this case, please consult with others who might remember more, such as parents or other relatives.*

A. Basic demographic information about you and your elders

	YOU	YOUR MOTHER	YOUR FATHER	ANY GRANDPARENTS WHO LIVED WITH YOU
Year of birth?				
Place of birth? (city, country)				
Where grew up? (city, country)				
Native language(s)?				
Other languages?				

1. Please list all the places you have lived since birth, including the approximate amount of time you spent there.

2. How do you self-identify? (e.g., white, Asian, Korean American, African American, hapa, gay, straight, working class, middle class, adoptee, etc., etc., etc.)

3. Your gender?

B. Early exposure to Korean and other languages

1. What was the primary language/dialect spoken at home when you were growing up? Were there any other languages spoken at home (if so, please list the languages spoken, and indicate who spoke them)?

2. If you encountered Korean with people (e.g., family members) whom you regularly saw while you were growing up, please fill in the table below with the relation of the Korean-speaking person and indicate how often you overheard this person speaking Korean (to others), how often this person spoke Korean to you, and how often you spoke Korean to this person by indicating **from what age to what age (months, years, etc.) you had this experience** in the box next to the appropriate frequency level (fill in more than one if the frequency of the experience changed as you grew older).

RELATION (e.g., dad, aunt, family friend, etc.)	*How often did you overhear him/her speaking Korean?*	
	FREQUENCY	AGE RANGE
	occasionally	
	about once/wk.	
	many times/wk.	
	once a day	
	all the time	
	occasionally	
	about once/wk.	
	many times/wk.	
	once a day	
	all the time	

RELATION, (e.g., dad, aunt, family friend, etc.)	*How often did s/he speak Korean to you?*	
	FREQUENCY	AGE RANGE
	occasionally	
	about once/wk.	
	many times/wk.	
	once a day	
	all the time	
	occasionally	
	about once/wk.	
	many times/wk.	
	once a day	
	all the time	

<table>
<tr><th rowspan="2">RELATION (e.g., dad, aunt, family friend, etc.)</th><th colspan="2">How often did you speak Korean to him/her?</th></tr>
<tr><th>FREQUENCY</th><th>AGE RANGE</th></tr>
<tr><td rowspan="5"></td><td>occasionally</td><td></td></tr>
<tr><td>about once/wk.</td><td></td></tr>
<tr><td>many times/wk.</td><td></td></tr>
<tr><td>once a day</td><td></td></tr>
<tr><td>all the time</td><td></td></tr>
<tr><td rowspan="5"></td><td>occasionally</td><td></td></tr>
<tr><td>about once/wk.</td><td></td></tr>
<tr><td>many times/wk.</td><td></td></tr>
<tr><td>once a day</td><td></td></tr>
<tr><td>all the time</td><td></td></tr>
</table>

If there are more people than will fit in the table, please describe in free prose how they contributed to your exposure to Korean.

C. Immediate and extended family

1. If your parents were not born in the U.S., how old were they when they came to the U.S.?

   Father: Mother:

2. What language(s)/dialect(s) do your parents speak to each other? If they use more than one, which language do they use with each other the most? If they speak to each other in a mix of languages, which languages do they mix?

3. What language(s)/dialect(s) do your parents speak to you currently?

   Father: Mother:

   If they use more than one, which language do they use with you the most?

   Father: Mother:

   If they speak to you in a mix of languages, which languages do they mix?

   Father: Mother:

4. What language(s)/dialect(s) do your parents speak to other people?

5. Language(s)/dialect(s) you speak to others currently:

	MOST OFTEN USED LANGUAGE	FREQUENCY OF USE				
		rarely	sometimes	half of the time	most of the time	all of the time
To father						
To mother						
To grandparents						
To siblings						
To friends						
To other people						

	OTHER LANGUAGE(S) USED	FREQUENCY OF USE				
		rarely	sometimes	half of the time	most of the time	all of the time
To father						
To mother						
To grandparents						
To siblings						
To friends						
To other people						

6. Do you have siblings (if so, how many and older/younger)? How often do your siblings speak Korean to your parents? How often do your siblings speak Korean to you?

7. Have you ever lived with your grandparents (if so, at what age and for how long)? How often do/did they speak Korean to your parents? How often do/did they speak Korean to you or your siblings?

D. Education

1. Have you taken any formal Korean classes before (including community/church school)?

   If yes:

   - What level were the classes?
   - Where were they held?
   - How old were you when you began taking them?
   - How often did they meet? (hours/week, weeks/year, etc.)
   - How long did you take them for? (years, months, etc.)

2. Not counting this trip, have you ever visited Korea before? (If so, where, when, and for how long?)

3. If you've ever lived in Korea, did you receive any formal education here? (If so, what was the highest grade you completed here?)

E. Language proficiency

1. Including English, what would you say your best language/dialect is?

2. When someone speaks Korean in a formal situation (e.g., speech, sermon), how much do you estimate you understand? (0%–100%)

3. When you watch TV in Korean without subtitles, how much do you estimate you understand? (0%–100%)

4. Please indicate your speaking ability in Korean in the following situations ('x' the appropriate box).

	LEVEL OF SPEAKING					
	none	poor	fair	good	excellent	native
Telling children's stories						
Ordering food at a restaurant						
Shopping						
Conversing with relatives about casual topics at family gatherings						
Conversing with strangers at a community meeting						
Talking about school or work						
Discussing politics						
Giving a speech						
Being interviewed for a job						

5. Please list any languages or dialects you speak or have studied besides Korean and English, and indicate the number of years you have spoken or studied it.

F. Pre-Korea

1. Before landing in Korea, how much time had you spent studying *hangeul*/Korean? (hours/week, etc.)

2. On a scale of **1 to 10**, how comfortable do you feel with Korean currently? (1 = not comfortable at all, 10 = couldn't be more comfortable)

3. On a scale of **1 to 10**, how comfortable do you feel with *hangeul* currently? (1 = not comfortable at all, 10 = couldn't be more comfortable)

4. On a scale of **1 to 10**, how motivated are you to learn Korean during orientation?
   (1 = not motivated at all, 10 = couldn't be more motivated)

5. On a scale of **1 to 10**, how motivated are you to become fluent in Korean?
   (1 = not motivated at all, 10 = couldn't be more motivated)

6. What is your ultimate goal with respect to learning Korean? Are you just trying to learn enough to get by at school? Is your goal to become fluent? Feel free to provide as much detail as necessary.

If you consulted with someone to fill out this questionnaire, please give their name and contact information:

Appendix B

Exit Questionnaire

Instructions: *The purpose of this questionnaire is to gather some information about your experience learning Korean over the past six weeks. Please fill it out as completely and accurately as possible.* ***All information will be kept strictly anonymous.***

A. Korean study and instruction

1. What Korean class were you in (e.g., Beginner A, Intermediate, etc.)?

2. How did each of your teachers either improve or detract from your learning of Korean? (If you do not know their name, please describe them clearly so that I know who you're talking about.)

3. On average, how much (e.g., %) did you understand of what was going on in your class?

4. On average, how much time per week did you spend studying Korean outside of class? (If the amount of time varied drastically from week to week, please indicate this—e.g., week #1 = 10 hours, week #2 = 20 hours, etc.)

B. Other exposure to Korean during orientation

1. Did you have a language partner during orientation? What was his/her name? About how much time per week did you spend interacting with him/her in Korean?

2. Were you involved in any other extracurricular activities in which you were exposed to Korean? If so, what were they, and how much time per week did you spend hearing/speaking Korean during these activities?

3. Did you ever engage the RAs in Korean? If so, how much time per week did you spend hearing/speaking Korean with them?

4. All things considered over the past six weeks (from class to activities to dorm life to going out), how much time per week do you estimate you were hearing Korean? How much time per week do you estimate you were speaking Korean?

C. Current Korean proficiency

1. When someone speaks Korean in a formal situation (e.g., speech, sermon), how much do you estimate you understand now? (0%–100%)

2. When you watch TV in Korean without subtitles, how much do you estimate you understand now? (0%–100%)

3. Please indicate your current speaking ability in Korean in the following situations ('x' the appropriate box).

	LEVEL OF SPEAKING					
	none	poor	fair	good	excellent	native
Telling children's stories						
Ordering food at a restaurant						
Shopping						
Conversing with relatives about casual topics at family gatherings						
Conversing with strangers at a community meeting						
Talking about school or work						
Discussing politics						
Giving a speech						
Being interviewed for a job						

4. On a scale of **1 to 10**, how comfortable do you feel with Korean currently? (1 = not comfortable at all, 10 = couldn't be more comfortable)

5. On a scale of **1 to 10**, how comfortable do you feel with *hangeul* currently? (1 = not comfortable at all, 10 = couldn't be more comfortable)

6. On a scale of **1 to 10**, how motivated are you now to become fluent in Korean? (1 = not motivated at all, 10 = couldn't be more motivated)

7. Have your goals with respect to learning Korean changed from the beginning of the summer? If so, what are they now?

D. Participation in the study

1. Do you have any history of speech, hearing, or language impairments?

2. What do you think the purpose of this study was?

3. In the listening test, did you follow any particular strategy in picking 바 apart from 빠 apart from 파? If so, please articulate what your strategy was.