Phonetic Convergence During Conversational Interaction in Bilingual Speakers

Emily Cibelli

Submitted in Partial Fulfillment
of the
Prerequisite for Honors
in Cognitive and Linguistic Sciences

April 2009

Table of Contents

Acknowledgments

First and foremost, I thank my family, whose support of my education and consistent encouragement has given me the confidence to pursue my research.

Of course, this work would not have happened without the guidance and knowledge of my advisors and committee members - Margery Lucas, Nancy Genero, Wendy Bauer, and especially my thesis and major advisor, Andrea Levitt.

Thanks are also due to the Office of the Dean of the College, who generously provided funding to offset my experimental costs.

Finally, I thank my friends for providing the inspiration for the topic of this research. Were it not for their constant (gentle!) heckling of my own tendency to pick up accents, I might never have found a topic that is so personally, as well as intellectually, gratifying.

Chapter 1 – Introduction to the Literature on Second-Language Phonetic Acquisition

Phonetic abilities in second-language acquisition

Second language acquisition is commonly considered a much more significant challenge to the language learner than first language acquisition. While all normally developing children who are exposed to language in the typical manner generally acquire their native language without a problem, many second language learners fail to achieve the same success. This difficulty is particularly common amongst adult second-language learners. Even those adult language learners who achieve a high degree of comprehension and fluency may have deficiencies in syntax, lexicon, or phonology that distinguish them from native speakers of the target language.

One attempt to explain this difficulty is summed up in the critical period hypothesis. The critical period hypothesis is a theory addressing the tendency in both first- and second-language acquisition for language learning to become significantly more challenging after a certain age, a trend that suggests that there is a "critical period" during which a person must acquire a language. The hypothesis has been proposed in varying degrees of strength by various researchers (see Hyltenstram and Abrahamsson 2003 for a review of the research), but most of them typically posit a drop-off in language acquisition abilities in late childhood or early teenage years. Even proponents of a weaker version of the theory, sometimes referred to as the sensitive period hypothesis (see Lamendella 1977 for the original use of the term, as well as investigations by Burrill 1985 and Hyltenstram and Abrahamsson 2000) suggest that adults will struggle with language acquisition more than children. Supporters of the critical period hypothesis predict that second language (L2) phonetic acquisition will present a particularly stubborn challenge to

adults, and that even second language learners who have acquired nativelike grammars and extensive vocabularies in their L2 can still be plagued by a foreign accent (Oyama 1976).

While many critical period proponents argue that language learners can learn a second language with nativelike abilities in their childhood, some researchers push back the cut-off for the critical period much earlier. Much of the literature on phonetic acquisition addresses the issue of whether or not a person's phonological inventory is fixed from infancy, and whether the phonetic system retains any plasticity after a very early age. One group of researchers (prominent among them Nuria Sebastián-Gallés) supports the notion that phonetic dominance in one language system emerges very early on in first language acquisition and is an absolute that does not change over a person's lifetime, despite exposure to and experience with other languages. This theory pushes back the critical period for phonetic acquisition to very early childhood. Even bilinguals who have acquired more than one language natively may be constrained by dominance effects – in other words, the claim is that there is no such thing as a truly "balanced" bilingual, and that every speaker will exhibit dominance in one phonetic system over another. Sebastián-Gallés and Soto-Faraco (1999) asked Catalan-Spanish bilinguals to choose which of two non-words presented on a screen corresponded to a non-word played out loud for them. The pairs of words were minimal pairs, with each pair containing either a vowel or consonant contrast that occurs only in Catalan, but not in Spanish. The authors showed that highly proficient bilinguals who learned Catalan as a second language at an early age still did not perform like native speakers of Catalan in the discrimination task.

Sebastián-Gallés et al. (2005) examined a similar phenomenon, by asking Catalan-Spanish bilinguals who spoke either Spanish or Catalan as their first language and who acquired their L2s at different ages to identify Catalan words and non-words. The non-words contained a

vowel contrast in Catalan that is difficult for Spanish speakers to hear, and the authors found that Spanish-Catalan bilinguals had trouble with this contrast even if they had learned Catalan at an early age. This study further demonstrated that even those speakers who were raised bilingually from birth exhibited evidence of having a dominant language; furthermore, these simultaneous bilinguals did not exhibit the same proficiency as monolingual native speakers of either one of their languages.

In a possible explanation for this trend, Bosch et al. (2000) argued that the dominance of L1 phonetic categories compromises the acquisition of phonetic categories in the L2. When asked to discriminate between three vowels (Catalan /e/ and /ɛ/ and Spanish /e/), participants could only discriminate the phonemes that occurred in their native language. These results, along with goodness of fit judgments on the vowel categories that the authors conducted, suggest that individuals were only sensitive to good exemplars of a given vowel category in their first language and not their second language.

Still, it should not be assumed that non-native *perceptual* abilities will necessarily prevent a second-language learner from achieving near-native *production* in their L2, especially if an individual receives extensive training and exposure to the target language. After all, it is possible that perception is more challenging than production, and that individuals who fail to perceive particular phonetic details will nevertheless be able to articulate them with a high degree of native-like ability. This idea is suggested by the work done in Bongaerts et al. (1995) (and reaffirmed by later work done in 2000), which demonstrated that L2 learners can achieve native-like production given the proper training and conditions that encouraged learning and use of the target language. Although the researchers did not examine the perceptual abilities of the participants in these studies, they demonstrated the important point that native-like production

can be attained as a second-language learner. Piske et al. (2002) asked native English speakers to rate the vowel production of speakers who had acquired English either early or late in life and who had maintained either a high or low continued usage of their L1. The results demonstrated that early bilinguals who showed low continued use of their L1 were rated as having L2 vowel production comparable to native speakers of the L2, indicating that the frequency of L1 use is directly tied to success in L2 acquisition. These studies suggest that native-like production can be obtained by second-language learners and that successful L2 production can achieve a high rating on a scale of native-like production by native speakers of the target language.

The issue of dominance and phonetic acquisition is really a question of plasticity, and various studies have attempted to examine how language history and experience affect phonetic plasticity. Many researchers (e.g. Thompson 2006, Piske et al. 2001, Flege et al. 1997) argue that maintenance of the first language inhibits nativelike production in the L2, which suggests that although an individual's phonetic capacity may be flexible, it is constrained by his or her *present* language usage, as well as his or her past exposure. Guion (2003) examined three groups of bilinguals – simultaneous bilinguals, early bilinguals, and bilinguals who acquired their second language later in life. All groups were asked to listen to recorded words containing particular vowels in Quichua and Spanish and then to repeat the words back. The author found that all groups were capable of phonetic plasticity in that their productions began to move in the direction of the recorded speech. However, simultaneous bilinguals were able to produce vowels that more closely matched the recorded speech than early bilinguals, suggesting an advantage in early native exposure (and indicating that the theory of absolute dominance in one language may not be true for production). Taken together, the results of these studies suggest that early language exposure has a significant effect on phonetic perception and production, but that many

other factors play a role in L2 abilities, and that the phonological system may be more plastic than would be suggested by strict critical period or dominance models. As mentioned earlier, phonetic plasticity may be particularly robust when it comes to production as opposed to perception.

The research done on L2 phonetic acquisition does not end at studies that probe the limits of individuals' current sound systems; many researchers have also investigated techniques designed to improve phonetic perception and production of the target language. This line of inquiry examines the conditions under which phonetic transfer and convergence could possibly occur. Many studies have shown that specific and targeted training can improve perception of L2 phonetic contrasts that do not occur in the native language (e.g. Cenoz and Lecumberri 1999, Callan et al. 2003). As noted earlier, Bongaerts et al. (2000) demonstrated that native-like production can be achieved by motivated L2 learners who are taught using a combination of input and instruction.

Interestingly, some researchers have suggested that exposure to the target language in an immersive environment is not the most efficient strategy for phonetic acquisition. Piske et al. (2001) recorded production of early and late L2 learners of English and asked native English speakers to rate how native-like the productions were. The study failed to find a correlation between length of residence in an L2-dominant country and native-like production in the L2. Bongaerts et al. (2000) are more optimistic about the effects of environmental exposure, noting that a combination of motivation to learn the L2 and exposure to an environment where the L2 was spoken produced speakers who rated as native-like speakers in L2 pronunciation, although they were on the low end of the range of native-like pronunciation and training was deemed necessary for participants to receive a high rating of native-like ability. From these results, we

can argue that although environmental exposure alone is not sufficient to produce a high degree of native-like production in the L2, it can nevertheless still have an impact on the phonological system of the learner if he or she is highly motivated.

Phonetic Convergence

Given the research discussed above, environmental exposure to a target language does not seem to strongly promote L2 phonetic acquisition. Still, it should not be ignored as a factor that can positively affect the phonetic production of the L2 learner. Learners can make a conscious effort to pick up the phonetic cues they are exposed to, but a largely unconscious process known as phonetic convergence also plays a part in acquisition. Pardo (2006) broadly defines phonetic convergence as "an increase in similarity among linguistic components." Within the specific aims of this paper, the term refers to an unconscious process by which listeners pick up phonetic cues from their conversational partners and from the ambient speech environment and incorporate them into their own phonetic production. Pickering and Garrod (2004) describe the process in their *interactive alignment account*, which proposes that conversational partners will align their speech automatically over the course of conversation (on many levels, including phonetically). This mechanism is not limited to bilinguals or second-language learners; in fact, it is commonly found between conversational partners who share a native language but who come from different dialect regions and thus speak with different accents (for an empirical investigation of this type of situation, see Delvaux and Soquet 2007). This tendency goes largely unnoticed in short interactions, but can be quite noticeable when a speaker is immersed in a particular speech environment for a prolonged period of time. Indeed, many speakers (the author included) who live in a new dialect region or a new country for enough time find themselves adopting the accent of the region, with no conscious effort or attempt to do so on their part.

Phonetic convergence has clear significance for second-language learning, as it ties directly to those effects on phonetic production that do exist as a result of environmental exposure. The phenomenon is also one potential demonstration of a link between phonetic perception and production, as individuals demonstrate their sensitivity to the speech they hear through their own speech production. It further highlights that phonetic learning as an adult is not *solely* a conscious process (although it is certainly facilitated greatly by training), as might be suggested by strong critical period hypotheses that posit that language acquisition skills drop off to insignificant levels after a certain age. It is a process that arguably affects all speakers immersed in a conversational context, although effect levels are likely to vary among individuals and populations.

Several studies have been devoted to examining the aspects of phonetic convergence in a number of circumstances, in part to account for differing levels of sensitivity to the effect in different populations. Sancier and Fowler (1997) did a case study of a Portuguese-English bilingual from Brazil who spent extended time in both Brazil and in the United States. The researchers measured the voice-onset time (VOT) – the length of time between the release of a consonant and the beginning of vocal cord vibration - of voiceless stops produced by the speaker in both languages, and found that these VOTs were much shorter after time in Brazil than time in the United States. This finding aligned with the relative VOTs of Portuguese and English and with the authors' predictions, indicating that the speaker was affected by the dominant language in each of her countries of residence. Of particular interest is the fact that the speaker's consonant production was affected in both languages by her stays in each country, even though she primarily spoke the dominant language of each country when in that location. In addition, the researchers measured native English and Portuguese speakers' perceptions of her production in

their respective languages. They found that only Portuguese speakers were sensitive to the VOT shift caused by the speaker's change of location; English speakers were insensitive to a shift. The authors attribute this difference between the populations to the fact that Portuguese speakers only had to distinguish between unaccented Portuguese (prior to the stay in the United States) and accented Portuguese (after the stay), whereas English speakers were required to recognize varying degrees of accented English, presumably a more challenging task.

Some researchers have tackled the issue of phonetic convergence by conducting empirical experiments designed to induce shifts in a short-term time frame. Flege and Eefting (1987c) ran such a study by manipulating the VOT of the phoneme /t/ in order to test the perception and production of native Dutch speakers who had varying levels of experience and proficiency in English. Previous research cited by the authors had shown that VOT values are shorter at category boundaries in Dutch than in English for the consonants /p/, /t/, and /k/, which are unaspirated in Dutch. As such, native Dutch speakers would be expected to have shorter VOTs, even when speaking English, than native English speakers. Participants were required to read material out loud in either English or Dutch (all were tested in both languages during separate sessions) and then identify the category boundary of a /da/-/ta/ continuum (a list of 16 recorded phonemes that transition incrementally from one phoneme to the other). Because the category boundaries of these phonemes differ between English and Dutch, the researchers were able to demonstrate that the language that the participants read in prior to making a boundary judgment affected their placement of the boundary. A post-test survey indicated that participants were generally unaware of this effect, and many incorrectly believed that the /da/-/ta/ stimuli differed between the two language conditions. However, the researchers found no significant difference in perception as an effect of English language proficiency, a result that stands in

contrast to the expectation that increased experience with a second language will lead to heightened perceptual abilities. Production data showed that many participants produced more English-like VOTs for the phoneme /t/ as a result of the stimuli and even suggested that their native Dutch production was affected by their L2 production in English.

Bail and Schmidt-Renfree (1999) set out to investigate how phonological priming in one language would affect production in a second language by conducting a study with two Danish-English bilinguals, one dominant in each language. Participants were asked to read a passage in one language and then to rapidly read off of flash cards written in both languages, with the expectation that the language of the passage would influence the participants' phonetic production when they read the flash cards. The effects found by the authors were limited – only one phoneme (/t/) showed carry-over effects that were distinct enough to be recorded in impressionistic transcription (although this method of analysis is highly suspect), and the two participants did not show very similar patterns. However, it is worth noting that they did find some evidence for the idea that L2 production can influence the dominant language, and not simply the other way around.

Pardo (2006) also elicited productions from participants that demonstrated short-term phonetic convergence in a study focused on more conversational production. Speakers of American English and British English either gave or received instructions and communicated with each other in order to complete a map task. Participants were paired in single-sex teams to avoid any sociological or hierarchical effects of gender on convergence. Researchers recorded participants' pronunciations of particular landmarks on the map before and after the task and compared these recordings to participants' production during the task. The study found that speakers converged upon one another's pronunciation over the course of the map task and that

the effects of this convergence carried over to the post-test recordings, when speakers were reading landmark names by themselves. The study also found effects of gender – surprisingly, to Pardo, than men converged more than women, despite her prediction that women would be more sympathetic conversationalists and thus more likely to converge. The role that participants played in the map task was also important - participants that gave instructions converged more to those receiving the isntructions than vice versa (although this pattern was dependent on gender). The results suggest that phonetic convergence may play a social role that goes above and beyond simple communication and transference of ideas.

Delvaux and Soquet (2007) looked for phonetic convergence in the context of a conversation between speakers of different dialects. They exposed speakers of two dialects of Belgian French to speech from the other dialect. They focused their study on vowels that differed between the two dialects but appeared in the same contexts. Participants were recorded during a test session "interacting" with a pre-recorded speaker from the other dialect region, as well as by themselves in comparison pre- and post-test conditions. To ensure that the process of phonetic convergence was unconscious, participants were shown pictures and asked to construct sentences based on them. The participants were told that this exercise was a memory task in order to direct their attention away from the accent of the speech they were listening to. The researchers found the effects of convergence during and after the recordings were played even though participants were generally unaware that the speaker had a different accent or that they had imitated them in any way. The authors described the unconscious effect as "mimesis" and contrasted this phenomenon with conscious imitation, which would be a deliberate attempt to imitate an accent. As in Pardo (2006), evidence for convergence in the post-test suggests that this phenomenon does continue to exert an influence for a least a short while, even after the participant is no

longer exposed to the accented speech. The authors ascribe this tendency to the fact that even though phonetic convergence is an unconscious process, it is robust enough to leave a memory trace and affect the speaker's productions even when he or she is not in the environment of the accent.

There are a number of broad conclusions to be drawn from the literature presented here. First, it is evident that the phonetic system of an individual's first language is strongly ingrained and that L2 phonetic acquisition can be quite challenging, possibly in part due to critical period or dominance effects. Secondly, although environmental effects are not striking on their own, the literature on phonetic convergence suggests that they do have a part to play in phonetic plasticity. It is possible that, contrary to intuition, unconscious phonetic acquisition may be easier than a conscious attempt to integrate the phonetic features of ambient speech into an individual's own production. Finally, it is worth noting that the effects of phonetic convergence seem to be most robust in studies examining vowel shifts as opposed to consonants, suggesting that some phonetic features may be more plastic than others.

This evidence takes us to the question this study will investigate: how will short-term priming for phonetic convergence affect speakers of differing levels of bilingual proficiency? In order to address this issue, I will be testing language users with varying language histories in English and Spanish. My work will depart somewhat from the model presented in many of the studies discussed above; instead of exposing speakers to speech as it is typically pronounced, I will instead investigate whether speakers' productions will phonetically converge with a model provided by foreign-accented speech. More specifically, I am interested in determining the extent to which bilingual speakers will imitate Spanish spoken with an American English accent and English spoken with a Spanish accent.

The vowel systems of Spanish and American English

Spanish and English are ideal languages for a comparison of bilingual speakers in this research. The population of Spanish-English bilinguals in the United States is quite significant, and this population is well represented at Wellesley College, the community from which participants are drawn for this study. Spanish-English bilinguals at Wellesley come from a number of different backgrounds: many have been raised in bilingual households in the United States, a number of others are originally from Spanish-speaking countries, and a third group have acquired Spanish as a second language and have come to study in the Wellesley Spanish department.

One could imagine dialectal variation as a potential concern given the wide variety of backgrounds that students come from. It is true that those recruited for the study may exhibit regional differences in the way they speak; however, in Spanish the more prominent phonetic variations are typically found in the consonants, whereas the research described here is concerned with vowel production. Although there is some vowel variation between the Spanish dialects, it is considered minor and much less prominent than consonantal variation (see Whitley 2002) and thus is not a serious concern for this study.

Another reason for selecting English and Spanish in this study is that their respective vowel systems have not only been widely studied but have also been extensively compared (see, as examples, Delattre 1965, Williams 1977, and Whitley 2002). These investigations have identified a number of similarities between the two systems but also some key differences, which are highlighted below. A few of these distinctions form the basis for the experimental manipulation, and these will also be discussed in detail.

Monophthongs

The inventories of monophthongs in Spanish and English are noticeably different. Whitley (2002) lists thirteen stressed and two unstressed vowels in the American English inventory, but just five in the Spanish phonetic system:

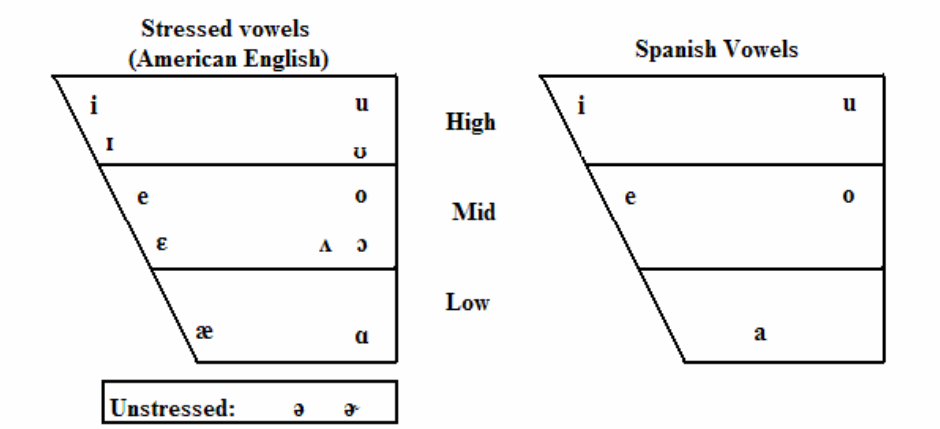
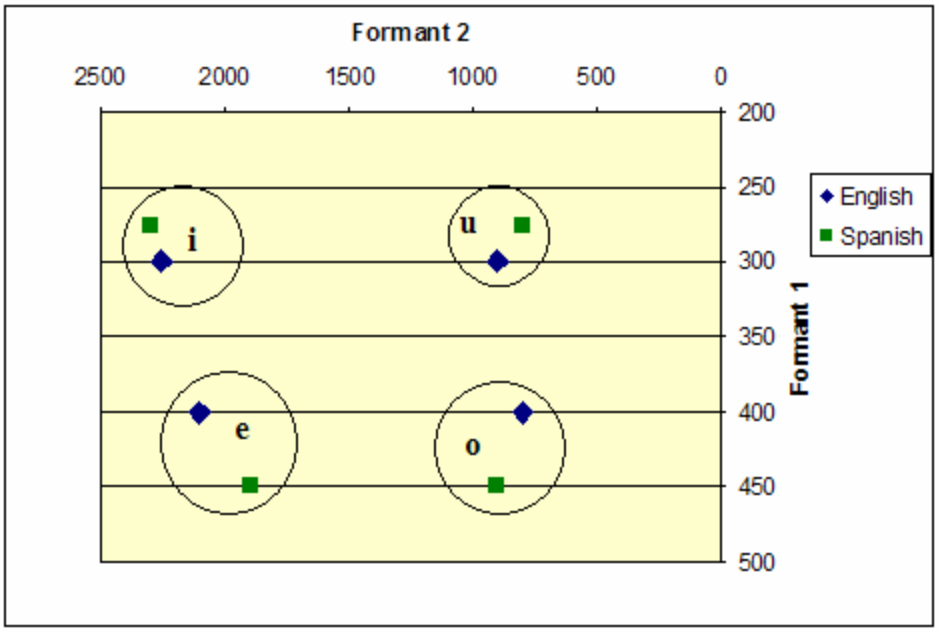


Figure 1. Vowel systems of American English and Spanish.

Note that the Spanish vowel set has no unstressed vowels – the five stressed vowels occur in both stressed and unstressed syllables and are not reduced as they would be in English (Martínez-Celdrán et al. 2003). As an example, the Spanish word "balón" would be transcribed as /balon/, retaining the /a/ in the unstressed syllable, whereas the English "balloon" would have a reduced vowel (ə) in the unstressed syllable (/bəlun/). The Spanish vowels are also "pure", meaning that the tongue holds a steady position throughout their articulation, while the vowels shared by English (/u/, /i/, and particularly /e/ and /o/), are typically diphthongized and are characterized by a change in the lip and tongue position as the vowel is articulated (Whitley 2002), even though they are represented by a single IPA character above.

The formant frequencies of the shared vowels in the two sets, although similar, differ slightly between the languages, as shown below:



Graph 1. Formant 1 and 2 of English and Spanish vowels /e/, /i/, /o/, and /u/.

(after Delattre, 1965[1]). Delattre also notes that vowel duration plays a part as a contrastive feature in English. The vowels /ɪ/, /ʊ/, /ɛ/, and /ʌ/ are significantly shorter than other English monophthongs, and their duration (along with other aspects of vowel quality) cause them to contrast with vowels such as /i/ and /u/. The Spanish vowel inventory, by contrast, does not have short vowels such as /ɪ/ and /ʊ/. However, it is true that stress causes vowel duration to play some role in meaning in Spanish words. While stress usually falls on the penultimate syllable in Spanish words, it can sometimes fall on other syllables, causing minimal pairs that differ only in the stressed syllable. These differences in stress do affect the duration of vowels in Spanish, but these duration differences are not considered to be the same as the distinction between English long and short vowels.

[1] There are two things to note when considering Delattre's data. One is that his measurements refers to an average male voice – a female voice would typically show formants 10-15% higher. The second is that the point in the vowel at which the measurement was taken was not indicated, a fact that is relevant because of the movement of formants in English vowels that takes place during diphthongization.

Diphthongs and Diphthongization

As mentioned above, American English speakers have a tendency to diphthongize the "pure vowels" that are shared with Spanish, particularly noticeable in the vowels /e/ (which is realized as /ei/) and /o/ (which is realized as /ou/). By contrast, those vowels in Spanish are pronounced with steady articulation rather than the slight diphthongization that would be found in American English.

It is worth noting that both languages also have an inventory of true diphthongs. Whitley (2002) lists four main diphthongs in English - three rising, moving from an open vowel to a glide with a higher tongue position, as well as one falling diphthong:

Rising	ai	ɔi	aʊ
Falling	ju		

Chart 1. Diphthongs of English, as described by Whitley (2002).

The Spanish inventory is broader; Martinéz-Celdrán et al. (2003) list 6 rising and 8 falling diphthongs:

Rising	je	ja	jo	ju	wi	we	wa	wo
Falling	e̯i	a̯i	o̯i	e̯u	a̯u	o̯u		

Chart 2. Diphthongs of Spanish, as described by Martinéz-Celdrán et al. (2003).[2]

The case for /ei/

One experimental manipulation in the current study will make use of a contrast between the Spanish /e/ and the English /ei/ (see the next section for a discussion of all phonetic contrasts

[2] It should be noted that transcription practices differ slightly based on various phonetic tradition, and that diphthongs or diphthongized vowels may have more than one standard transcription for the same sound (for example, /eɪ/ is frequently used in place of /ei/). The transcriptions in Chart 1 and Chart 2 reflect the IPA symbols chosen by the authors cited.

in the study). Because both the Spanish and English phonetic inventories contain the phonetic sequence /ei/, it may seem misguided to portray /e/ vs. /ei/ as a contrast between the two languages. However, the comparison is valid, as the two languages make use of the diphthong/diphthongized vowel in distinct environments and with different qualities. A first clue to the distinction is in the orthography; Spanish words that make use of /ei/ will typically be spelled as to indicate it (compare the /ei/of "reina" to the /e/ of "mesa"). Such an indication is not made in English orthography, as the English /ei/ is a vowel that has the quality of being diphthongized, rather than being a true diphthong. Hualde (2005) notes that the Spanish /e/ and /ei/ *do* contrast phonologically, whereas the American English /e/ is *always* realized as /ei/. San Francisco et al. (2006) found that amongst a group of Spanish-English bilinguals, those instructed in Spanish were more likely than those instructed in English to see diphthongs as two units due to Spanish phonology and orthography, a result that indicates that the English and Spanish /ei/ are perceived differently. Finally, Whitley (2002) notes that there are distinctions between the articulations of the diphthong in the two languages. The Spanish /ei/ is marked by a quick upglide and a high final position, while the English /ei/ "dwells" on the /e/ for a longer duration and does not rise as high as its Spanish counterpart. These distinctions all justify the choice to contrast the Spanish /e/ with the English production of /ei/, despite the fact that the /ei/ sequence can be found in both languages.

The experimental manipulations

This study will examine three particular phonetic manipulations that take advantage of the differences between the phonetic systems of Spanish and English. These contrasts were

chosen because the differences in the Spanish and English vocalic system lead me to believe that individuals may substitute the corresponding vowel or diphthongized vowel that is found in their native language for the vowel or diphthongized vowel in the other language, particularly if they are inexperienced in the second language. Furthermore, the recorded speech that will be played to participants will exhibit these features, so that even experienced speakers may be influenced to use the vowel of one language while speaking the other language. The comparisons being used are:

Spanish	**Sample word**	**English**	**Sample word**
/e/	mesa	/ei/	table
/e/	teléfono	/ɛ/	telephone
/o/	bote	/ou/	Boat

Chart 3. Phonetic comparisons and sample words for critical vowels of the current study.

In addition, preliminary test recordings I have conducted indicate that vowel durations differ between native Spanish speakers and native English speakers, even when both are speaking the same language. As a result, this study will also look for differences in vowel duration by different speakers.

The tendencies described above provide an opportunity to examine the effects of phonetic convergence in speakers of Spanish and English. Recordings made by late second-language learners in their second language will reflect the vowel systems of their native language, even though the words they speak are in the second language. As an example, a native English speaker who has learned Spanish late in life may pronounce the Spanish word /mesa/ (*mesa,* table) with an English pronunciation and render the word as /meisa/ (or perhaps even /meisə/, although the final schwa is not of primary interest for this particular study.)

Similarly, a native Spanish speaker who has learned English late in life may pronounce the word "boat" as /bot/, rather than the English /bout/. Such pronunciations reflect a disconnect between the language and the vocalic system being used. By exposing participants to these utterances, I will explore the extent to which the foreign-accented speech is mimicked, which will indicate the strength of the unconscious phonetic convergence tendency.

Chapter 2: Experimental Design, Procedure, and Predictions

Participant Populations

Four populations of speakers with different language backgrounds were recruited for the current study. The first group is comprised of "balanced" Spanish-English bilinguals – those who have been raised bilingually from birth, or from a very early age, and express equal levels of fluency in their two languages. These will be hereafter referred to as the balanced bilinguals. The second population contains native English speakers who have learned Spanish as a second language after late childhood or early adolescence. These speakers will be fluent in both languages, but clearly dominant in English. They will be referred to as the English-dominant bilinguals. The third population is the reverse – native Spanish speakers who have acquired English as a second language – and will be referred to as the Spanish-dominant bilinguals. These three populations were tested in two sessions, one exposing them to English-accented Spanish, and a second to Spanish-accented English. Sessions were counterbalanced so that equal numbers of participants were exposed to the English and the Spanish session first. Finally, a group of monolingual native English speakers (hereafter referred to as the monolingual group) acted as controls and were only exposed to Spanish-accented English. This group provided a baseline for comparison by demonstrating how much phonetic convergence (if any) can be obtained when speakers are exposed to an accent that is foreign to the language they speak.

Participants were recruited from the Wellesley College student body. Nineteen female participants, aged 19-23, participated in the study. Six recruited participants were balanced bilinguals, five were English-dominant bilinguals, three were Spanish-dominant bilinguals, and five were English monolinguals. One participant in the balanced bilingual category was excluded from the analysis because she was unable to complete the second session of the study.

Participants were compensated $10 (English monolinguals) or $20 (all bilingual populations) for their participation in the study.

Stimuli recordings

Two speakers were recruited to make the recordings used as stimuli in the experiment. The individuals were chosen because aspects of their speech exhibited the desired non-native speech qualities for the study (for example, diphthongized vowels produced while speaking Spanish.) The speaker of the English-accented Spanish sentences was a third-year female Wellesley College student who had studied Spanish in high school and during one summer semester in college. She also speaks Chinese at home. The speaker of the Spanish-accented English sentences was a male Wellesley College employee who was born and raised in Puerto Rico. He studied English for one hour a day while in school in Puerto Rico, but did not achieve fluency until the age of 24, five years after starting college in the United States. He also spent time studying Japanese.[3]

For both recording sessions, the speakers were asked to repeat each critical sentence three times. The best example of each sentence was selected by evaluating the degree of accent in the vowels, as well as with consideration for clarity and consistency (tokens that were recorded with a creaky voice quality or were not produced fluently were eliminated). Because of ambient noise in the background during the English-accented Spanish recordings, the sound stimuli in this condition were also cleaned up using the program Audacity[4] to remove extraneous background noise.

[3] Fortuitously, neither of the stimuli speakers' third languages pose a problem for the experimental predictions. Chinese has a number of diphthongs, and so the speech of the English-accented Spanish speaker had the influence of two languages that make use of diphthongs. Conversely, Japanese does not have diphthongs, and so the speech of the Spanish-accented English speaker was influenced by two languages with relatively monophthongized vowels.
[4] Audacity Team (2009): Audacity (Version 1.3.4-beta) [Computer program]. Retrieved April 4, 2009, from http://audacity.sourceforge.net/.

Experimental Design and Procedure

The experimental design for this study is based on the work of Delvaux and Soquet (2007), which examined phonetic convergence by exposing speakers of two dialects of Belgian French – the Brussels and Liège dialects – to speech recorded by a speaker from the dialect they did not speak. In general, the experimental procedure described below replicates that of the original study (with the exception of some details, such as the number of critical vowels used in each study.) However, a few major changes were made to suit the hypotheses and participants available for the current study. For one, the comparisons being made in the current study are comparisons between languages, not dialects. Secondly, the participants in the current study do not come from backgrounds of identical proficiency – some are second-language learners of one or other of the languages, and some our bilingual while others are monolingual. Finally, while the participants of the original study were speakers of Belgian French, the current study participants are speakers of Spanish and American English.

As mentioned above, all bilingual participants (balanced bilinguals, English-dominant, and Spanish-dominant) were tested in two sessions, one exposing them to English-accented Spanish and another to Spanish-accented English. Sessions were counterbalanced between the two language conditions. Because English monolingual participants had no experience with Spanish, they were only tested in one session and exposed only to Spanish-accented English.

At the beginning of their first (or only) session, all participants took a language background survey (see Appendix A for the full survey). The survey includes questions about a participant's acquisition of his or her language(s), educational information, demographic information, and the participant's self-assessment of his or her speaking and listening proficiency in each of the languages. It was based on a recent study undertaken by Lim et al. (2008), who

developed a survey as a tool to assess language dominance. The researchers found that amongst a group of Mandarin-English bilinguals, the survey was more successful at sorting participants into Mandarin-dominant, English-dominant, and balanced bilingual groups than measures such as age of first exposure and years of exposure. Because of its reliability in predicting language dominance, I chose to use the survey (in a modified form) in the present study to document a more detailed history of each participant's language background. Any anomalies or irregular patterns in the data, as well as any interesting individual trends, may in part be explained by the information gathered from the survey.

The experimental portion of the study was split up into five parts – a training section, a pre-test section, two test sections, and a post-test section. In all sections, participants were exposed to PowerPoint slides like the one below:

Figure 2. Sample slide from the visual stimuli used in the experiment.

Each of these slides contained three things: a number, an object noun (in the above, the telephone), and a location noun (in the above, the table.) For each slide, participants constructed pre-formulated sentences to incorporate the three elements in the slide, using this pattern:

"There is/are [number] [object noun(s)] on the [location noun]."

"Hay [number] [object noun(s)] en el/la [location noun]."

In the example above, the sentences constructed would be "There are two telephones on the table" and "Hay dos teléfonos en la mesa" for the English and Spanish sessions, respectively. (For a complete list of sentences and images used, see Appendix B.) In the training session, participants were exposed to slides to familiarize themselves with the images, and were given a guide sheet to help them practice the sentence construction. In the other four sections, they constructed sentences without the guide.

All sections of the experiment were run similarly, with a series of slides that prompted the participant to either construct a sentence or to listen to a pre-recorded sentence. The pre-test section required the speaker simply to construct a sentence for each slide, which provided the baseline phonetic characteristics of the speaker. The test 1 and test 2 sections required a combination of speaking and listening. These sections exposed the participant to the accented speech and served as the auditory stimuli that participants could converge to in their own production. The post-test section was again the participant speaking alone, without any auditory stimuli. The breakdown of sentences for each section was as follows:

Section	Total number of slides	Sentences spoken by participant	Sentences spoken by pre-recorded speaker
Training	27	27	0
Pre-test	54	54	0
Test 1	162	54	108
Test 2	162	54	108
Post-test	54	54	0

Chart 4. Number of slides presented in each section of the experiment.

To signal to the speaker whether he or she would speak or listen on a given slide during the test sections, the following indicators were given:

Figure 3: Sample slide with auditory stimuli

Figure 4: Sample slide with prompt to speak

In the slide on the left, the participant would listen to a pre-recorded sentence that played automatically when the slide was reached. The instruction to listen is also indicated by the speaker symbol on the bottom of the slide. In the slide on the right, there is no embedded sound, and the arrow at the bottom of the slide instructs the participant to speak by constructing a pre-formulated sentence according to the format described above.

The unit of analysis for all sections of the experiment was the vowels produced by the participant in their speech. The images used in the study were carefully chosen because they contain a critical vowel in the stressed syllable. The critical vowels, discussed in the first chapter, are the Spanish vowels /e/ and /o/ and the English vowels /ɛ/, and /ei/, and /ou/. Each object noun and each location noun contains a critical vowel, so that in each sentence participants either spoke or listened to two critical vowels. The nouns, their critical vowels, and the corresponding vowels of the other language are as follows:

Spanish nouns		
Words	**Spanish vowel**	**Corresponding English Vowel**
banqu**e**ta	e	ei
bol**e**ta	e	ɛ
f**o**to	o	ou
p**e**sa	e	ei
r**o**sa	o	ou
tel**é**fono	e	ɛ
banqu**e**ta	e	ei
b**o**te	o	ou
m**e**sa	e	ɛ
English Nouns		
Words	**English vowel**	**Corresponding Spanish Vowel**
ph**o**to	ou	o
(sheet of) p**a**per	ei	e
r**o**se	ou	o
sh**e**ll	ɛ	e
t**e**lephone	ɛ	e
w**ei**ght	ei	e
b**oa**t	ou	o
shelf	ɛ	e
table	ei	e

Chart 5. List of critical words and their corresponding vowels.

The corresponding vowels refer to the vowels produced by the pre-recorded speakers in the speech stimuli, and accordingly also correspond to the predicted phonetic shift that participants would exhibit in their speech after being exposed to the accented speech stimuli. For example, the Spanish word “bote” contains the vowel /o/, however, it was pronounced with the diphthong /ou/ by the pre-recorded speaker, rendering it /boute/. Thus, the corresponding English vowel for this word is /oʊ/. In the case of Spanish words containing /e/, which theoretically could shift to either /ei/ or /ɛ/, the predictions were made based on cognate

equivalents in English (telephone/teléfono)[5], a translated equivalent that contains one vowel or another (weight/pesa), and the production of the speaker who recorded the Spanish sentences. The distribution of vowels throughout the words and sentences were equalized so that a participant heard each of the critical vowels an equal number of times. (The exception to this is the Spanish /e/, which occurred twice as frequently as /o/. However, because the words containing /e/ were predicted to shift to two different English vowels, this was counted as an even distribution.)

All sections of the test were recorded, which allowed me to compare participants' production before, during, and after exposure to the recorded sentences they were exposed to in the test section. All sections of each session were recorded on an external microphone using the voice-activated software Express Dictate[6]. A back-up microphone also recorded data onto a CompactFlash card. Recording sessions took place in the recording studio of the Knapp Center in Clapp Library at Wellesley College.

Because I am interested in eliciting an unconscious effect, participants were not told to match the phonetic production of the speaker. Instead, I followed Delvaux and Soquet (2007) and instructed participants that their job was to complete a memory task, and that they should focus on the words in the sentences they are speaking that appear on the screen. To test their memories, participants were asked questions about the location of objects, the colors of numbers they saw, and the frequency of particular images. These questions were asked after the test 1 section of the experiment in both sessions, and at the end of the first (or only) session. This

[5] In some pairs, such as telephone/teléfono, the critical vowel fell on different (stressed) syllables. However, the effect of cognate equivalence was assumed to carry over from one syllable to the other.
[6] By NCH Software. http://www.nch.com.au/express/index.html.

controlled for any *intentional* efforts by participants to imitate the speech they are hearing by focusing their attention on an alternate task.

Following the completion of the final (or only) experimental session, all participants were asked to complete a post-test phonetic assessment that evaluated their attitudes towards their accent production in each language. (See Appendix C for the full assessment.) Participants were asked to judge how native-like they believed their accent to be in English and Spanish (or English only, for monolingual speakers). They were then asked if they had any awareness of imitating others' speech, a measure intended to indicate the degree to which individuals are aware of their own phonetic convergence. This survey will help to tease apart trends by potentially demonstrating relationships between accent self-perception, awareness of phonetic convergence, and performance on the experimental task.

Phonetic Analysis

I analyzed two aspects of each vowel using the phonetic analysis software Praat[7] in order to determine the degree of convergence over the course of the experiment: vowel duration and formant frequencies. Formants are the frequencies at which a vowel resonates, and they can be used to distinguish one vowel from another. The movement of the first, second, and third formants over the duration of a vowel can also indicate the diphthongization of a vowel as it transitions from one sound to another:

[7] Boersma, Paul & Weenink, David (2009). Praat: doing phonetics by computer (Version 5.0.46) [Computer program]. Retrieved January 7, 2009, from http://www.praat.org/.

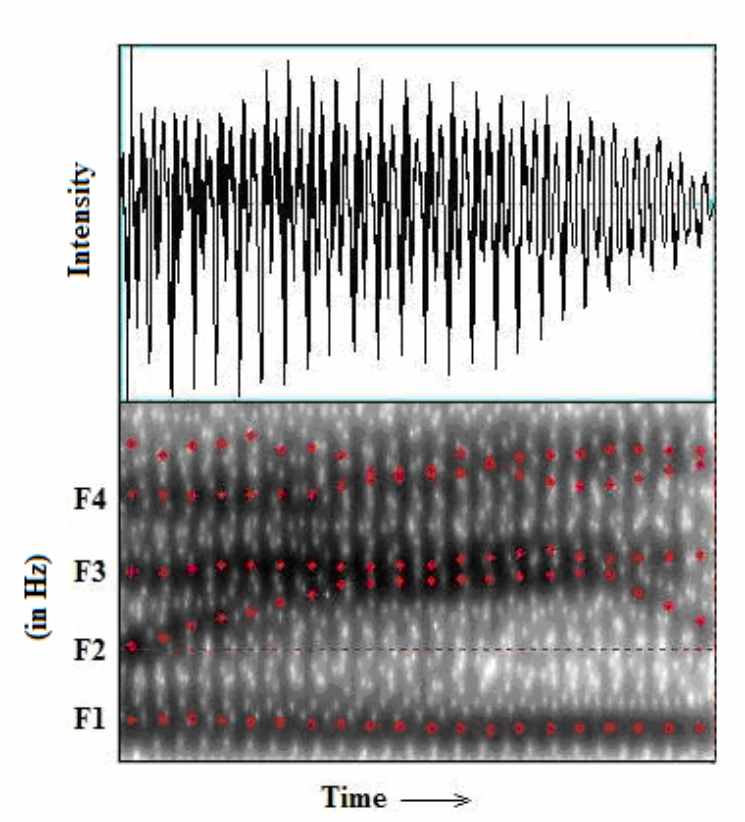


Figure 5. Sample diphthong: /ei/ from the English word "weight". Red dots represent the formants in the vowel over the course of time.

The above image shows the waveform (top) and spectrogram (bottom) of the vowel /ei/ in the English word "weight". The dark dotted lines in the spectrogram indicate the formants. Notice in particular the upward glide of the second formant as the vowel transitions from the /e/ to the /i/.

The duration analysis comprised a simple measure of the length of each critical vowel. The beginning and end of each vowel was identified by examining both the waveform and spectrogram of the sound for the characteristic qualities of the vowels (signaled by a sustained sound with voicing, significant intensity, and particular formant patterns). The beginning of each vowel was marked after the transition from the preceding consonant, and the end marked before the transition to the following consonant, in order to avoid coarticulation effects that would affect the formants of the vowel.

The formant analysis was more involved: the first, second, and third formants were measured at the one-third point in the duration of the vowel, and then again at the highest point of each formant. The difference between the high point and the one-third point of each formant was calculated to get the overall formant change for each vowel, which represented the amount of movement corresponding to diphthongization (or a lack thereof) in the vowel. These changes were plotted over the course of the experiment to determine whether a speaker's tendency to diphthongize or monophthongize a vowel, as well as to alter the duration of the vowel, was influenced by the speech he or she was exposed to during the course of the experiment. Both length and formant measurements were made manually in Praat.

Over the course of the experiment, bilingual speakers produced 432 critical vowels and monolingual speakers produced 216 critical vowels. Due to the time constraints of this project, the time it would take to manually analyze these vowels would be prohibitive. As a result, only a subset of vowels were included for analysis. In each of the four sections of each experimental section, six examples of each critical vowel were analyzed, two from each carrier word. The two examples of each word came from the first utterance of the word in the section and the last utterance of the word in the section. This sample allowed me to track the change in vowel production over the course of each section, as well as the changes between sections.

Experimental Predictions

If the results of the present study are comparable to those of Delvaux and Soquet (2007), all participants should show most sensitivity to the recorded speech during the test phases, and show a slightly lessened shift during the post-test phase, although this effect will still be heightened over the baseline of the pre-test phase. Across groups, there are a number of possible patterns that could arise as a result of differences between the populations or individuals. The

relationships between language history and degree of phonetic convergence have been largely unexamined. Because of this lack of previous research specifically addressing the issue and the largely exploratory nature of this work, I am not committing to a specific prediction, but instead will consider a range of potential outcomes for the data.

Broadly speaking, there are three possible outcomes for the data as a whole. It is possible that each individual will react differently to the stimuli presented, and each show his or her own pattern of convergence that does not necessarily conform to the population group he or she is in. Alternately, it may be that individuals across all groups show a similar tendency towards convergence, suggesting that the phenomenon is a universal tendency that is not affected by proficiency or experience with a language. Finally, it is possible that groups as a whole will exhibit different trends, with some showing more sensitivity to phonetic convergence than others. If that is in fact the case, some potential outcomes for each group are as follows:

I argue that balanced bilinguals could demonstrate any of a number of patterns of production, and the results derived from this group will be particularly telling. Sancier and Fowler (1997) demonstrated that language users are adept at recognizing the effects of phonetic convergence (i.e. a shift in a speaker's accent) in the speech of others when those speakers share the listeners' native language. Conversely, when a listener's first language is spoken by the speaker as a second language, the shift in the speaker's accent in that language is much more difficult for the listener to perceive. Since speakers in the balanced bilingual population are native speakers of both languages being presented, they should be very conscious of the foreign accents in both sets of recordings. This heightened awareness may lead this population to resist phonetic convergence, as they would be converging upon an accent that was "foreign" and thus violated their native phonological inventory. If this is the case, we can expect the same thing

from the English monolingual speakers, and indeed the effect of phonetic convergence exhibited by these two groups may prove to be similar.

However, as Sebastián-Gallés et al. (2005) have noted, simultaneous early bilinguals do not perform identically to native monolingual speakers of their languages on phonetics-related tasks, and the fact that balanced bilinguals have two phonetic systems may suggest that they will have a lessened sensitivity as compared to monolingual speakers, even if this difference is slight. Thus, a demonstration of phonetic convergence amongst balanced bilinguals, particularly if it contrasts with the monolingual controls, will carry weighty implications for the abilities of early bilinguals as opposed to monolingual speakers and may lend support to the idea that bilinguals are not simply "two monolinguals in one body." (For a discussion of this issue, see Grosjean 1989.)

The patterns of results shown by the English-dominant and Spanish-dominant bilinguals will have important consequences for the literature on second-language acquisition. It is conceivable that these participants will show the biggest effect of any group when they are tested in their second language, and there are two reasons to suspect this. In the first place, they did not acquire this phonetic set natively, unlike the monolinguals and balanced bilinguals, and so they may be more susceptible to foreign accented speech in their second language, as well as less attuned to detecting it and rejecting it as foreign. Secondly, the "foreign accent" will come from their native language, and share the phonetic characteristics of their native phonetic system, and so they may be even more inclined to imitation as a result of their increased experience with this particular phonetic inventory.

However, it will also be interesting to see how these participants react to accented speech presented in their native language. Past and recent work has suggested that acquisition of an L2

phonetic system has an effect on L1 phonetic production, particularly when the speaker attains fluency in the second language (e.g. Williams 1979, Flege and Eefting 1987c, Jiang 2008). This suggests that the phonetic systems of the bilingual are dynamic systems which can be influenced by one another, even when one has obvious "dominance" over another. Sancier and Fowler (1997) also demonstrated that a native Brazilian's Portuguese was influenced by prolonged stays in the United States and surrounded by English, the speaker's L2. So if the current study finds English-dominant and Spanish-dominant bilinguals demonstrating a meaningful effect in their native language as well as their second language, this tendency may suggest that participants do not identify a "foreign" accent as being a language-phonetic mismatch as readily as would a monolingual speaker. Furthermore, it would lend support to the idea that even the dominant native phonetic system is susceptible to modification over the course of an individual's lifetime, based on their experiences and input.

Monolingual English speakers may show the most resistance to phonetic convergence. The model for this study, Delvaux and Soquet (2007), tested individuals from different dialect regions and found that their production showed evidence of phonetic convergence even though many were unaware that they were listening to a speaker with a different accent. However, although participants in that study exhibited regional differences, they were all native speakers of Belgian French and the input they received was from a "valid" variation of their native language. Monolingual speakers in this study, by contrast, will be listening to a foreign accent (Spanish-accented English) and are more likely to recognize the speech as coming from a non-native speaker, given the incongruence between language spoken and phonetic inventory. In other words, their awareness of the distinctions between a native American accent and a Spanish-accented speech sample will cause them to reject the influences of the accented speech, because

that accent is marked as foreign by their native experience with English. If the monolingual speakers do show resistance to phonetic convergence as a result of this incongruence, it would suggest that listeners are less sensitive to accents that are influenced by experience with a different language, as opposed to an accent from a dialect that is a subset of the speaker's own native language. A foreign accent may present a strong enough incongruence between language and phonetic system for the monolingual participants to unconsciously reject any significant effects of phonetic convergence. Of course, if the monolingual English speakers do prove to be sensitive to Spanish-accented English speech, it will suggest that phonetic convergence is an incredibly robust phenomenon and is pervasive enough to affect even those speakers who recognize that the accent they are listening to is non-native for the language being spoken.

The above discussion has progressed with the assumption that the different populations being tested will exhibit different patterns of sensitivity to phonetic convergence. However, a lack of significant difference between the populations will also have some interesting implications. Although I would not be able to make claims about the nature of first and second language acquisition if no between-group effects were found, such results would still say something about the pervasiveness of phonetic convergence. In particular, if all populations across both languages show significant convergence, these results will suggest that all individuals are sympathetic to phonetic cues from the external environment and that language history is not the mitigating factor in how language users acquire new phonetic information. Alternately, if no significant results are found across groups, but some individuals show a phonetic shift, it will indicate that the phenomenon of phonetic convergence is highly idiosyncratic, and that individual variation prevails over any group tendencies.

Chapter 3: Experimental Results

Statistical Analysis

The primary statistical analyses for the production data were a series of repeated measures ANOVA run in SPSS. Eight separate ANOVAs were run – four for the English-accented Spanish data (duration, formant 1 change, formant 2 change, and formant 3 change), and four for the Spanish-accented English (analyzing the same four measurements). The reasoning for running separate analyses was twofold. First, the small sample sizes meant that the statistical power of the data was already greatly reduced, and running one large analysis with more factors would obscure the already small potential to see effects and trends in the data. Second, the comparisons that were separated out into different analyses were not deemed to be relevant to the hypotheses. In particular, comparing duration and formants 1, 2, and 3 is not a particularly useful measure of the trends of convergence, as they are largely separate measures. Comparing the Spanish and English sessions would have potentially interesting consequences, but preliminary investigation of the data made it clear that the trends in both sessions were not strong enough to merit direct comparison. Thus, the eight analyses run were on the following measures:

1. Spanish-accented English (SAE) Duration
2. SAE Formant 1 Change
3. SAE Formant 2 Change
4. SAE Formant 3 Change
5. English-accented Spanish (EAS) Duration
6. EAS Formant 1 Change
7. EAS Formant 2 Change
8. EAS Formant 3 Change

The "formant change" measure refers to the calculated difference in formant measurements between the high point of the formant and the 1/3 point of the formant that was discussed in Chapter 2.

One shortcoming of the phonetic analysis methodology I used was that it did not account for the pattern of formant 1 movement in the American English /ei/ and /ou/. In these vowels, the movement from /e/ to /i/, and /o/ to /u/, reflects a decrease in formant 1, but the analysis I used only captured upwards movement of the formants. As a result, the statistical analyses of formant 1 (numbers 2 and 6 in the list above) were unreliable, and they are excluded from this report. Formant 2 and formant 3 are predicted to rise over the course of diphthongization in these vowels, so the analysis on these two formants was justified. (For precise measures of the average formants of these vowels, see Delattre 1965, Ohala 1994, and Hillenbrand et al. 1995).

Each statistical analysis involved repeated measures ANOVA with two within-group factors and one between-group factor. The first within-group factor, "time," measured whether vowel production varied over the four experimental sections (pre-test, test 1, test 2, and post-test). The second within-group factor, "vowel," measured whether vowel production varied between the specific critical vowels (e/ei, o/oʊ, and e/ɛ). The between-groups factor, "population," measured whether vowel production varied between the four populations of participants.

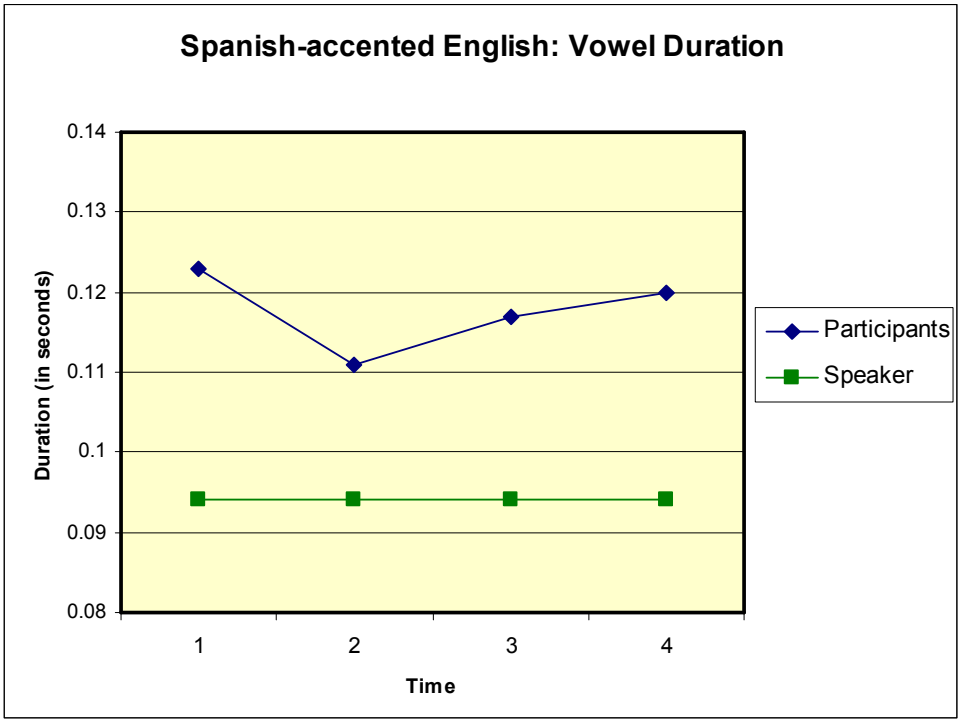
The Effect of Time

The "time" variable was used to measure the extent to which participants converged upon the vowel production of the speaker at each of these four times. This effect demonstrated not only whether or not convergence occurred at all, but also whether the effect of convergence increased with increased exposure (seen by a strengthening of convergence at time 3), and whether lingering effects of convergence led to a continually shifted production after participants were no longer exposed to accented speech (time 4).

Spanish-Accented English Data

Duration analysis

The data on vowel durations in the Spanish-accented English session show the clearest pattern of evidence that participant productions were affected by speaker input. There was a main effect for time [$F (2.173, 30.427) = 7.332$, $p = .002$, $\eta_p^2 = .344$].



Graph 2. Main effect of time for vowel duration, Spanish-accented English session.

As can be seen from the chart above, there is a clear effect of the exposure to accented speech on vowel duration. The blue line represents the average vowel duration, collapsed across vowel and participant group, at each of the four times in the experimental session. The green line represents the average vowel duration of the stimuli speaker's productions. The pattern above demonstrates that participants were affected by exposure to accented speech at time 2, which caused their vowel durations to shorten in a convergence to the speech being played to them. Bonferroni post-

hoc comparisons show that the decrease in duration from time 1 to time 2 is significant ($p = .001$, adjusted alpha = .0125). The increase in duration at time 3 ($p = .052$), although not significant, suggests that participants may have become aware of the change in their productions, and moderated the convergence effect to an extent. However, even when participants spoke by themselves again at time 4, the effect of convergence on duration has not been completely lost, as durations still at time 4 still sit below the starting durations at time 1. Although this comparison was not significant, the general trend of the data suggests that this interpretation holds some validity, and could be replicated with larger sample sizes.

Formant analysis

Formant analysis was less conclusive with respect to time. There was no significant main effect for time in the analysis of formant 2. This result indicates that the movement in formant 2 of speakers' vowel productions did not significantly change as a result of exposure to the Spanish-accented English speech they listened to in the test 1 and test 2 sections of the experiment.

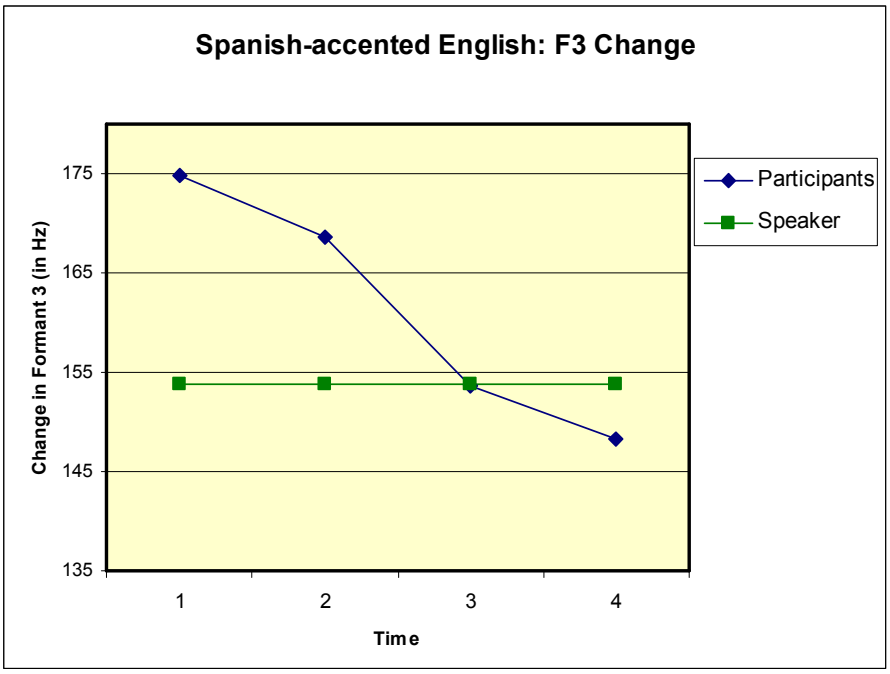
One potential confounding factor in the formant analysis was the inclusion of the e/ɛ vowel in the analysis with the other vowels. Because neither /ɛ/(the vowel in one-third of the critical stimuli words) or /e/ (the Spanish vowel that was used in the accented speech for those particular words) is a diphthong, there is no reason to expect that we would see movement in formant 2 in either native English or native Spanish speaker production during the Spanish-accented English session in those particular words[8]. To examine this potential confound, another repeated measures ANOVA was run excluding the e/ɛ data, and examining only the data of

[8] The same is not necessarily true of the Spanish-accented English session. I predicted a shift from /e/ to /ɛ/ in this condition, but it is possible that some speakers would shift their /e/ production to /ei/ instead, and as such, movement in formant 2 would still be a relevant measure in this case.

e/ei and o/ou production. However, the main effect of time on formant 2 movement was still not significant when the e/ɛ data was excluded. Thus, the exclusion of e/ɛ data did not impact the overall lack of effect of time on formant 2 movement.

The results of formant 3 analysis were more promising than formant 1 or formant 2. There was a significant main effect for time [$F (2.306, 32.279) = 3.498$, $p = .036$, $\eta_p^2 = .200$]. This is somewhat surprising, as formant 3 movement is a less striking measure of diphthongization than formant 2 movement. However, there is still some movement in the third formant corresponding to the shift from one vowel to another in diphthongization[9]. This result indicates that exposure to Spanish-accented English did affect participants' production of formant 3; more specifically, the movement in formant 3 production was lessened as a result of exposure to the accented speech. There was also a significant linear trend of time for formant 3 change [$F (1) = 5.804$, $p = .030$, $\eta_p^2 = .293$].

[9] For discussion and examples of spectrograms depicting formant movement in diphthongs, see Carmel, T. Spectral Cues for English Phonemes. In *Spectrogram Reading*. Retrieved April 18, 2009, from http://speech.bme.ogi.edu/tutordemos/SpectrogramReading/ipa/ipahome.html.



Graph 3. Main effect of time for formant 3 change, Spanish-accented English session.

The graph above illustrates the fact that formant 3 movement decreased linearly with exposure to the accented speech. Interestingly, formant 3 change at time 4 dipped below even the speaker's average F3 change, indicated that the effect of the accented speech continued to have a strong and increasing effect on participants' vowel production even after they were no longer being exposed to the speech. This pattern does not correspond with the duration data above, which showed an initial dip in duration at time 2, followed by a climb back to the "baseline" duration of time 1 at times 3 and 4. The difference between the two patterns may suggest that duration change is dependent on continued exposure, and may even be a salient enough difference that listeners will reject incorporating it into their speech upon repeated exposure. On the other hand, formant change, at least in the case of formant 3, seems to have been more readily adopted by participants, suggesting that the effects of monophthongization or diphthongization on production may continue to be an effective cue even after a listener is no longer exposed to the influencing speech. Post-hoc comparisons found no significant differences between time 1, 2, 3,

or 4. Thus, no conclusive statements can be made about the change from one specific section of the experiment to another with respect to formant 3 change. However, the significant linear trend is enough to establish that participants were adjusting their production of formant 3 with heightened sensitivity to the speaker at each time. This indicates that there was some moderating effect of the Spanish-accented English exposure such that participants monophthongized their vowels in response to exposure to the accented speech.

English-Accented Spanish Data

Duration and formant analysis

Unlike the data from the Spanish-accented English session, the analysis of the English-accented Spanish productions showed no main effect for time. This result indicates that participants were not adjusting their vowel durations to match the English-accented Spanish data they heard. Bonferroni post-hoc comparisons indicated that this failure to converge occurred across all population groups. The formant analysis of the English-accented Spanish session showed a similar lack of convergence. There was no significant effect of time on formant 1, formant 2, or formant 3. The lack of convergence in participants' speech as a result of exposure to the accented speech means that participants were not influenced to diphthongize their Spanish vowels to adopt a more English-like production.

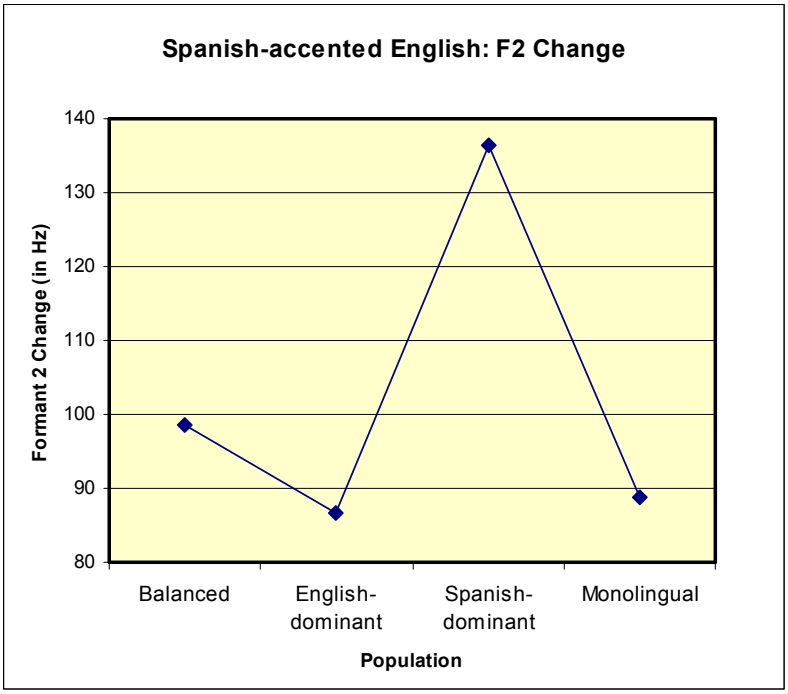
The data presented above indicates that, when participants converged upon the speech produced by the accented speaker they heard, they only did so in the Spanish-accented English session. By contrast, participants resisted convergence in the English-accented Spanish session. The disparity between the two sessions suggests that participants were either less aware of the Spanish accent of the English speaker, or else considered the English-accented Spanish a more egregious violation than the Spanish-accented English, and resisted efforts to converge in the

former session more than the latter. This somewhat surprising contrast will be discussed in detail in the Discussion in the next chapter.

The Effect of Population

Spanish-accented English data

Overall, there was no significant main effect of population for either vowel duration or formants in the Spanish-accented English sessions. However, the data for formant 2 production approached significance (p = .063).



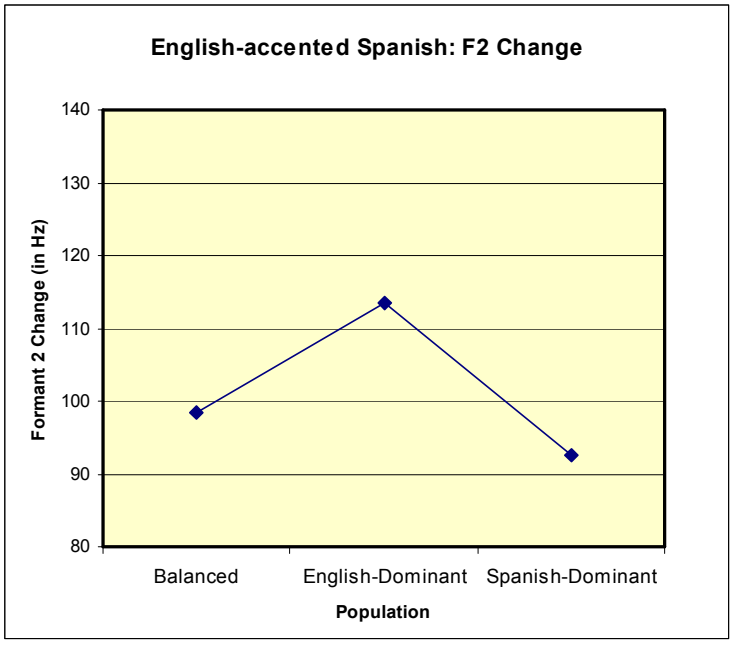
Graph 4. Main effect of population for formant 2 change, Spanish-accented English session.

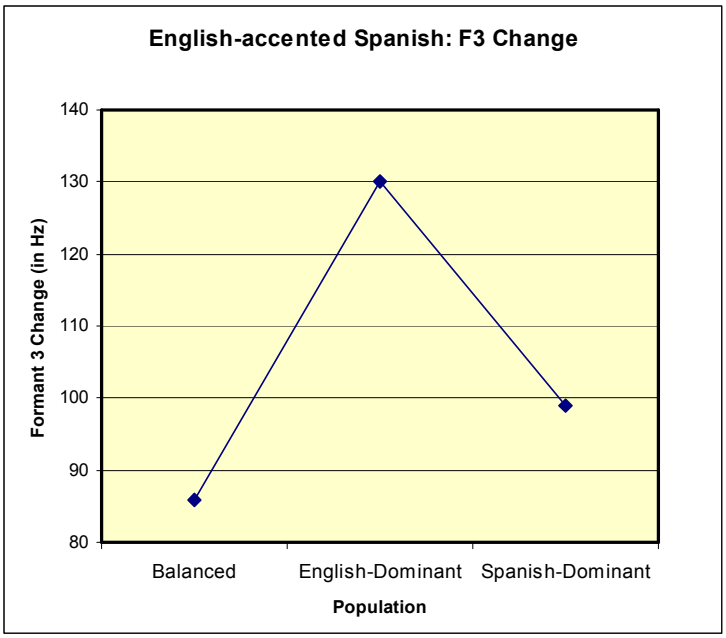
It is clear from the graph that Spanish-dominant bilinguals (population 3) are showing higher diphthongization in the second formant (a greater F2 change) than the other population groups. Although the post-hoc comparisons did not show significant differences between any of the groups, there were trends in the comparisons of the Spanish-dominant bilinguals to the English-dominant bilinguals (p = .086) and, to a lesser extent, the English monolinguals (p =

.110). This surprising result suggests that Spanish-dominant bilinguals may have been resisting efforts to converge to the Spanish-accented speech more than other populations by diphthongizing their vowels, a result that will be explored more thoroughly in the Discussion section.

English-accented Spanish data

While there was no significant main effect of population in the English-accented Spanish data in either duration or formants, the formant analysis showed patterns that are suggestive of group trends. The data plots for formants 2 and 3 demonstrated that English-dominant bilinguals showed more diphthongization in their vowels (a higher value for F2 and F3 Change) as compared to both the balanced bilinguals and Spanish-dominant bilinguals:





Graphs 4 and 5. Main effect of population for formant 2 and formant 3 change, English-accented Spanish session.

None of the above main effects were significant, but the main effect for population in formant 2 approached significance ($p = .087$).

The results presented above do not provide information about the participants' convergence to the accented speaker, as there was no "population x time" interaction and thus no indication that one population was converging more at a particular section of the experiment than another. However, the data do indicate overall how the speech of the populations compares to one another. The distinction between diphthongization in the English-dominant (population 2) and Spanish-dominant bilinguals (population 3) is unsurprising, as it reflects the expected prediction that English-native speakers would diphthongize their Spanish vowels, whereas Spanish-native speakers would use monophthongs while speaking Spanish. However, the graphs above suggest that the balanced bilinguals (population 1) were more Spanish-like in their formant production, as their F2, and F3 change values sat closer to those of the Spanish-dominant bilinguals than those of the English-dominant bilinguals.

The Effect of Separate Vowels

As mentioned above, the ANOVA analysis included "vowel" as a within-group factor, in order to test whether participants or population groups would be more likely to converge upon one vowel over another. Although there were significant differences between vowel durations and formants for all three vowels across the analyses run, there were no interactions between "vowel" and either "time" or "population." Thus, the analysis of separate vowels did not speak to my hypotheses, as it did not show differences between population groups or differences in convergence at different times in the experiment.

Post-Test Assessment

At the end of the last session of testing, participants were asked to fill out a post-test phonetic assessment evaluating their perception of their phonetic capabilities and habits. The assessment first asked participants to rate how native-like they felt their accents were in English and Spanish, respectively. Specifically, the questions asked:

1. How do you judge your accent to be in English?
2. How do your judge your accent to be in Spanish?

Participants rated their accents on a scale from 1 – 7, where 1 represented the idea, "I speak English/Spanish with a very heavy foreign accent", and 7 represented the idea, "I speak English/Spanish like a native speaker." Participants were then asked two questions regarding their perceptions about the extent to which they phonetically converge with other conversational participants:

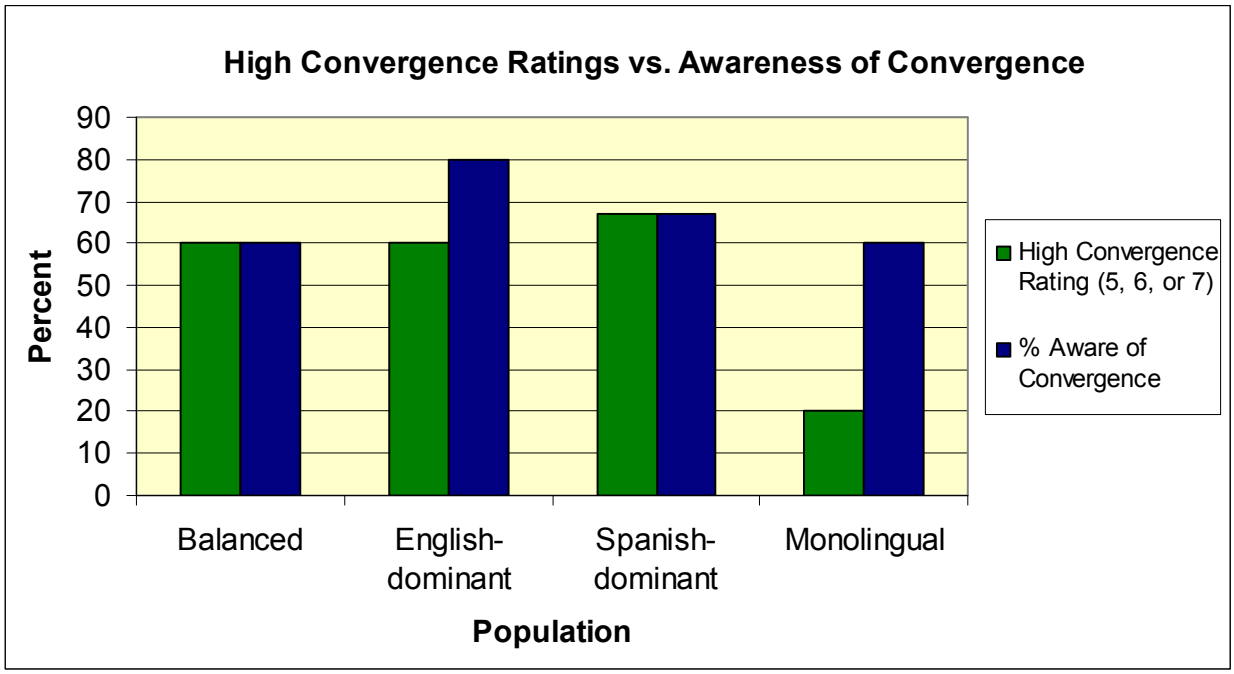
3. How much do you imitate the speech sounds of others, intentionally or otherwise? (Scale from 1 [never] to 7 [always])
4. Are you aware when this happens? (Yes/No)

Perception/Performance Connections

To test whether self-perception of accent related to performance during the experiment, correlations were run to compare responses to question 1 and 2 to production during test 1 and test 2 sections of the experiment of duration and formants. Although there were some significant correlations, they were scattered throughout different measures and different vowels, and do not paint a cohesive picture of correlation. Because nothing conclusive can be determined from the data, I have excluded them from analysis here; however, a more thorough examination of the correlations can be found in Appendix D.

Perceptions of Convergence

I was also interested in comparing whether the convergence tendencies referred to in question 3 were related to self-perception of convergence that was asked about in question 4. The results of the convergence questions are represented in the chart below, split into population groups. The green checkered bars represent the percentage of participants who answered question 3 with a 5, 6, or 7 (termed a "high convergence rating"), and the blue striped bars represent the percentage of participants who responded "Yes" to question 4.



Graph 6. Percent of responses of "5, 6, or 7" on question 3 and "yes" responses on question 4.

A simple correlation was run to determine whether there is a correlation between the number of participants who rated themselves as having a high convergence rating and the number of participants who were aware of converging. The test demonstrated that there was no significant correlation between the two measures. However, the graph above suggests that there may be a pattern related to bilingualism, where bilingual participants (populations 1, 2, and 3) rated themselves with higher convergence ratings than monolingual participants (population 4).

Chapter 4: Discussion and Conclusions

Discussion

The most conclusive result to come out of the analysis is the main effect of time for duration and movement in formant 3 (F3 change) in the Spanish-accented English condition. These effects indicate that participants' productions converged to the Spanish-accented English they heard in this session. Specifically, vowel durations dropped and vowels monophthongized in the formant 3 as a result of exposure to the accented speech. These results indicate that when exposed to Spanish-accented speech, participants as a whole were sensitive to the accent and (consciously or unconsciously) manipulated their own production to converge with the speech they heard. This result replicates in part the effects of convergence found by previous work. In particular, as Delvaux and Soquet (2007) demonstrated, the results show that convergence is at least partly an automatic process. Participants had no social or conversational impetus to converge to the accented speech in the experimental setting of this study, as they were not engaging in conversation with another individual. Therefore, there must be some semi-automatic process that causes some measure of convergence in production even when there is no conversationally cooperative reason to do so.

The pattern of results in the Spanish-accented English between the duration changes and the formant 3 movement are worth considering together, as their patterns of convergence differ over the course of time. Participants' convergence on the duration measure was most striking at time 2 (the test 1 section), suggesting that participants were responding quickly to the vowel durations of the accented speech when they first heard it. At time 3 (the test 2 section) the durations climbed back up again, suggesting that duration was no longer a strong indicator of convergence. However, formant 3 change in participants' production at time 3 continue to drop

and matched the accented speaker's level at time 3, which was an increase in convergence from time 2. At time 4 (the post-test second), when durations continued to climb again, increasing divergence on that measure, formant 3 change continued to drop, demonstrating more convergence on that measure.

One way to reconcile the patterns of these two measures is to consider that convergence may be a dynamic process, and that its effect is manifest in different measures at different points in exposure to speech. In the data presented here, the strongest trend of convergence is found in duration at time 2 and in formant 3 at time 3. This pattern reflects a shift in the feature that most strongly converges to the accented speech as the time spent "interacting" with the speech progresses. Arguably, participants may have adjusted their speech based on repeated exposure to the accented speech in an effort to modify their relationship with the speaker. Although there is no way to discern the extent to which this switch was conscious in the minds of participants, it nevertheless seems possible that their awareness shifted at some point between time 2 and time 3, which resulted in the transfer of strong convergence from duration to formant 3 movement.

Spanish-accented English vs. English-accented Spanish

One important difference in the overall production of participants was the convergence to accented speech in the Spanish-accented English condition, but not the English-accented Spanish condition. Although there was not a direct statistical comparison of the two sessions, the difference between the results in the two is still striking. Participants converged in duration and movement in the formant 3 when listening to English spoken with a Spanish accent, but showed no significant effects of convergence when listening to Spanish spoken with an English accent. This distinction is somewhat surprising. Given that all participants are immersed in an English-speaking environment while in the college setting, I predicted that English spoken with a Spanish

accent would be a more salient language-phonetic mismatch, and participants would be more likely to resist convergence when exposed to this speech. However, the pattern of results showed the opposite trend.

One way to make sense of this result is to rethink how individuals might respond to speech in their second language spoken with the accent of their first language. As described in the Results section, Spanish-dominant bilinguals behaved somewhat unlike other populations in their formant 2 production when presented with Spanish-accented English. In contrast to other populations, they showed more diphthongization, suggesting that they resisted convergence to the relatively monophthongized speech of the Spanish-accented speaker they were listening to. This trend may indicate a desire by native Spanish speakers to distance themselves from an accent that they might also produce, in an effort to sound more native-like in English, their second language.

If this theory is correct, it may help to explain why the results failed to show convergence in any measure for participants when exposed to English-accented Spanish. English-dominant bilinguals, all of whom are currently or have recently studied Spanish as a second language, may have a particularly strong motivation to avoid convergence to their native accent when speaking Spanish, as they are currently invested in learning Spanish with more native-like production. Similarly, balanced bilinguals may have resisted convergence in an effort to demonstrate that their Spanish production was equally as native as their English production. Spanish-dominant bilinguals would have been most "clued in" to the English accent, as Spanish is their native language, and it is not hard to see why they would avoid convergence in such an environment.

One other aspect of the stimuli may have affected the difference between the results of the English-accented Spanish and Spanish-accented English sessions – the fact that the stimuli of

the former were recorded by a female speaker, and the latter by a male speaker. If gender played a part, it would lead to the interesting conclusion that female speakers would be more likely to converge to accented speech spoken by a speaker of the opposite gender than a speaker of the same gender. One possibility is that the gender of the male Spanish-accented English speaker caused female speakers to focus on the qualities of his speech relating to his gender and made them less attuned to his accent, which may have permitted more unconscious convergence to his recordings than to the recordings of his female counterpart. It seems unlikely that this was the only component contributing to the differences between the results of the two sessions, but it is possible that it was a contributing factor.

Population effects

Although there were no significant effects of population on production in either the English-accented Spanish or Spanish-accented English conditions, the data did show some promising trends of population differences that require comment. The most salient trend was the near-significant effect of population on formant 2 movement (F2 change) in the Spanish-accented English condition. Graphing the average F2 change of the different populations in this condition demonstrated that Spanish-dominant bilinguals showed a strongly higher F2 change value – indicating more diphthongization – than participants in other populations. Because the accented speech was Spanish-accented, leading to the prediction that participants' productions would become monophthongized with exposure to the speech, this result is particularly notable because it indicates strong divergence from the Spanish-accented English on the part of the Spanish-dominant bilinguals.

This trend is surprising, given the prediction that participants listening to speech in their second language that has the accent of their first language would be less likely to detect the

accent and more likely to converge to the accent, as it uses the phonetic set of their native language. However, the trend of the data supports an alternative hypothesis that participants who are confronted with accented speech in their second language will strongly diverge from the speech in an effort to sound less like the accented speech and more native-like in the production of their second language. This interpretation conforms nicely with my anecdotal experience with Spanish-dominant participants; more than once during the experiment, I was asked if their accents mattered and whether they should try to suppress them for the recordings. This self-awareness indicates that Spanish-dominant participants were mindful of their accents throughout the English session, and may have felt a desire to diverge from the Spanish-accented speech they heard in an effort to distinguish their own speech from "foreign"-accented speech.

The mirror result – hyper-monophthongization by English-dominant participants in the English-accented Spanish condition – was not achieved. This asymmetry is logical, however, given the English-dominant environment of the college. English-dominant bilinguals may not be confronted very often about their American accents, as they spend most of their days in an American English-speaking environment. Even in their Spanish-language classes, their American English-accented Spanish pronunciation may not be corrected often, and they may have many other classmates who share their American English accent. Spanish-dominant bilinguals, by contrast, are in a language environment where their Spanish-accented English would be noticeable in the majority of situations, as native Spanish speakers are in the minority in most environments at the college. This does not mean that Spanish-dominant bilinguals spend all of their time at the college worrying about their English pronunciation; however, it is not surprising that they might become aware of it in an experimental setting where they are being recorded and are simultaneously listening to Spanish-accented English produced by another

native Spanish speaker. Thus, it makes sense that Spanish-dominant bilinguals would hyper-diphthongize their English vowels in an effort to diverge from the recorded speaker, even where English-dominant bilinguals would not hyper-monophthongize their Spanish vowels.

The results in the preceding chapter also suggested that when exposed to English-accented Spanish, balanced bilinguals' formant production was more similar to Spanish-dominant bilinguals than English-dominant bilinguals. Specifically, their formant change values for formants 1, 2, and 3 were comparatively low, indicating less diphthongization than English-dominant bilinguals and more monophthongization, like the production of the Spanish-dominant bilinguals. To ensure that this pattern was not a consequence of demographic factors, I reviewed the biographical information of the participants in the balanced bilingual population. 60% of the participants in this population were born in the United States, while 40% were born in Spanish-speaking countries in North and South America. When asked to rate their understanding and speaking skills in Spanish and English, 60% of the population ranked their English skills as the language they comprehend and speak the best, with Spanish ranked as the second best. 40%, by contrast, ranked their English and Spanish comprehension and speaking skills as equally good.

The biographical information of this population does not paint the picture of balanced bilinguals with a Spanish bias – indeed, if anything, the population seems to be slightly English-biased. Therefore, their tendency to produce vowels with Spanish formant structures more similar to Spanish-dominant bilinguals is unlikely to be due to a Spanish-dominant bias in their language history. It is likely, however, that their native or native-like experience with Spanish made the American English accent of the Spanish speaker they listened to more salient than it was for English-dominant bilinguals. Recognition of a foreign accent in the speech may have led balanced bilinguals to resist convergence to this speech. As such, the production of balanced

bilinguals looked more like the production of the Spanish-dominant bilinguals, who would also be aware of the accented speech, even if the balanced bilinguals did not come in with a strong Spanish bias in their language skills. This interpretation must be considered with caution, as the differences between the populations in production were not significant. However, the trends in the data do suggest that this interpretation may hold true upon further investigation.

Correlations

As mentioned in the results, the correlations I ran between post-test assessment questions and performance during the experiment did not result in a meaningful pattern of results. Although nothing conclusive can be drawn from the correlations between self-perception of accents in English and Spanish and test performance, there are some potential avenues for further investigation, which are discussed in Appendix D.

Although there was no significant correlation between awareness of phonetic convergence and a high convergence rating (see graph 7 in the Results section), the pattern shown in the data suggests that there may be a difference between bilingual and monolingual participants with respect to self-perception of convergence. Approximately 62% of bilingual participants self-reported a high convergence rating (5, 6, or 7), whereas only 20% of monolingual participants did. This suggests that self-perception of convergence may be related to bilingualism, and raises an interesting "chicken or egg" question. One the one hand, bilinguals may be more likely than monolinguals to exhibit high convergence tendencies because of their experience with multiple phonetic inventories and need to maintain more than one accent to use with the languages they speak. On the other hand, it may also be the case that those who have a higher propensity for convergence are the individuals who are more likely to

become successful bilinguals, as they show the adaptability and flexibility necessary to master a second accent.

Methodological improvements for future work

The implementation of this study was met with some methodological challenges, and future research that addresses these issues will help demonstrate more robust results and more confidence in the hypotheses under investigation. The most obvious challenge to this research in the college setting that it was undertaken is an asymmetric language environment. Wellesley College is an institution that operates in English, and American English is the predominant language spoken in the surrounding area. Thus, all participants in the study use English on a daily basis, and are required to be fluent in the language to attend the college. While they are at college, English will be the most common language of day-to-day life, as it is the shared language of the community, and so most participants at the time of testing were immersed in a predominantly English environment. This creates an asymmetry with respect to Spanish usage, as participants are not exposed to Spanish (and particularly not a native Spanish environment) nearly as much as English, nor do they have the opportunity to speak it as much in the college setting. Thus, both production and perception will be affected by increased exposure to English. For a more equal picture of language usage, a complementary study should be undertaken in a predominantly Spanish-speaking environment, to provide a basis of comparison.

Another problem of environment caused difficulties for participant recruitment. Although the Spanish-English bilingual population is one of the largest bilingual populations at Wellesley, it was still difficult to find speakers who fit the criteria for participation. This was particularly evident in the Spanish-dominant bilingual category, as I was only able to recruit three participants who were appropriate for the study from this population. A larger university setting,

or a location that has a more substantial bilingual population, would provide for larger population numbers. This is critical for testing the hypotheses I have laid out, as larger populations are necessary in order to see significant between-group differences in production.

A third potential area of improvement for future work is the quality of sound stimuli used in the experiment. Although every effort was made to find speakers for the stimuli recordings who exhibited the desired accent with respect to vowel production, the features of their accents were in many cases more subtle than was desirable. This was particularly the case with formant structure – the Spanish-accented English stimuli still showed a fair amount of diphthongization, and the English-accented Spanish vowels were often not far from monophthongs in their formant movement. This meant that the target accent, on the features that were analyzed, were often not very strong, providing at best a weak model for convergence. (Duration differences were generally more robust.) Future work would benefit from stimuli with more distinct accents.

Another potential confounding variable in the phonetic analysis was the consonant environment surrounding the critical vowels. Because coarticulation effects can affect the vowels immediately preceding or following a consonant (see Hardcastle and Hewlett 1999 for a review), accurate measures of the formant movement as a result of vowel articulation, as opposed to coarticulation effects, were often hard to achieve. Stimuli that better controlled for consonant environment would provide a clearer set of vowels for analysis.

Finally, the analysis of similar studies would benefit from semi-automated Praat procedures. For the current study, all phonetic analysis was done by hand. Due to the extremely large volume of available data, only a small subset of the critical vowels produced could be analyzed. This led to larger statistical margins of error, as only a few exemplars of each vowel comprised the mean for each participant's duration and formant measures. A semi-automated

procedure – auto-segmentation of critical vowels and automatic duration measurements, for example – would alleviate the amount of work needing to be done by hand and allow more data to be analyzed, which would provide larger sample sets that would increase the statistical power of the analyses.

Conclusions

I have undertaken the work presented above in order to determine the extent to which individuals will alter their production to match recorded speech in a laboratory setting. The work presented here has provided some support to the idea that individuals will show some phonetic convergence towards speech they are listening to even in the absence of a conversational context. This result suggests that phonetic convergence may be a semi-automatic process, operating even when there is no cooperative or social impetus to do so.

The results of this study also suggest that the extent that an individual's productions exhibit the effects of phonetic convergence may be controlled or monitored by the individual. This can occur because an individual becomes aware of his or her shift in production and may wish to temper it for various reasons, including maintenance of individual identity through his or her accent and avoidance of adopting a non-native accent[10]. This self-monitoring may be affected by the degree of the accent of the speech that the individual is listening to, as well as the relation that he or she has to that accent. Specifically, individuals who are second-language learners of the language being spoken may avoid converging with another second-language learner's accent in an attempt to sound more native-like in the second language.

[10] In natural conversation, another social factor that speakers may take into consideration if they are aware of phonetic convergence is respect for the individuals who they are speaking to. Speakers may wish to avoid offending their conversational partners by avoiding phonetic convergence, which could be seen as malicious mockery or imitation instead of an unconcscious process, particularly if the conversational partner has a strong accent. However, this factor was not examined in this study, as participants did not engage in natural conversation with the speakers they listened to.

This experiment was motivated by an interest in discovering whether the extent of phonetic convergence was affected by a person's language history – specifically, whether different types of bilingualism affected the phonetic plasticity of an individual's speech in the domain of convergence. Unfortunately, small sample sizes prevented any potential population differences from reaching significance in the current study. The results do suggest that speakers' language history may affect their *conscious* decisions about whether to converge or diverge based on the language they are speaking, their native language, and their relationship to the language history of the conversational partner or speaker of the model speech they are listening to. However, the current study cannot conclusively demonstrate that differences in *unconscious* cognitive mechanisms that stem from a language history of bilingualism or monolingualism might affect speakers' patterns of phonetic convergence.

Future Directions

The experiment described here lays the groundwork for future investigation into the question of phonetic plasticity and the role that language history plays in phonetic convergence. Beyond the methodological improvements discussed above, there are a number of ways that new research can explore and pursue the questions asked by this study. One way to investigate this issue would be to examine languages with more diverse phonetic inventories. Although the phonetic inventories of Spanish and English are not identical, they share a large number of phonemes. Their similarities may aid in conversational convergence, as speakers may make use of their similar features to facilitate their adaptation to their conversational partners. Future work could focus on second-language learners who are attempting to acquire a language with a phonetic set that is very different than that of their native language, in order to determine

whether convergence is affected when the phonetic features of the languages being spoken and the accents of speakers are very dissimilar.

Developmental work may shed light on phonetic plasticity and convergence from a different perspective. Comparative studies on phonetic convergence run with children and adults could suggest important things about the plasticity of phonetic abilities over the lifespan and contribute to the discussion on the critical/sensitive period. In addition, this work would shed light on the amount that cognitive capabilities and social factors play a role in convergence. It may be that children will demonstrate more conversational convergence than adults, as their phonetic systems are typically regarded as more plastic than older speakers. However, it is also possible that adults may be more sensitive to the social demands of conversation, and thus more likely to converge in an attempt to be more cooperative with their conversational partners.

The discussion above alluded to gender differences as a potential factor playing into the differences between convergence in the English-accented Spanish and Spanish-accented English conditions. As these differences were potentially confounded with other variables (including the language being spoken and the dominant language of the testing environment), it is impossible to ascertain the exact effect that gender may have had in the current study. Future work would benefit from an investigation of gender as another factor in this paradigm, as well as studies done in more naturalistic settings. Pardo (2006) avoided mixed-gender pairs in her conversational experiment in order to control for the effects of social hierarchy, but it may be worth examining those cross-gender effects (as well as the effects of single-sex pairs) in a bilingual experiment. This type of work will shed more light on the social factors that contribute to convergence, both in naturalistic (conversational) settings and in more controlled environments such as the one used in the current study.

Finally, a way to bring this work into a broader sphere is to look at convergence in features other than vowels, and even outside the domain of phonetics. Some preliminary work has been done on consonants and voice-onset time (e.g. Flege and Eefting 1987), prosody (Coulston, Oviatt, and Darves 2002), and speech rate and conversational turn duration (Street 1984). Expansion of these studies to include a bilingual component will allow us to compare various features of conversational convergence. This will shed light on what features are most salient in conversation, what types of linguistic details are most susceptible to convergence, and what aspects of speech may be more or less plastic depending on a speaker's language history. Other potential measures of investigation include, but are not limited to, lexical choice and terminology selection, syntactic structure, and paralinguistic features such as gesture.

Research on phonetic convergence is an important line of investigation because it speaks to both the social functions of language and the cognitive capabilities that underlie the linguistic system. My experiment is a preliminary exploration of the ways that language history and cognitive plasticity contribute to the social function of communication. Future work that builds on and extends the findings of this research will contribute to a fuller picture of phonetic capabilities, conversational strategies, and the nature of bilingualism.

References

Bail, M. J. and Schmidt-Renfree, N. (1999). Phonological Priming and Phonetic Carry-Over in Two Danish-English Bilinguals. *Linguistica Atlantica,* 21, 1-12.

Bongaerts, T., Planken, B., and Schils, E. (1995). Can Late Starters Attain a Native Accent in A Foreign Language? A Test of the Critical Period Hypothesis. In D. M. Singleton and Z. Lengyel (Eds.), *The Age Factor in Second Language Acquisition: A Critical Look at the Critical Period Hypothesis* (pp. 30-50). Multilingual Matters.

Bongaerts, T., Mennen, S., and van der Slik, F. (2002). Authenticity of Pronunciation in Naturalistic Second Language Acquisition: The Case of Very advanced Late Learners of Dutch as a Second Language. *Studia Linguistica,* 54(2), 298-308.

Bosch, L., Costa, A., and Sebastián-Gallés, N. First and Second Language Vowel Perception in Early Bilinguals. *Cognitive Psychology* 12(2), 189-221.

Burrill, C. (1985). The Sensitive Period Hypothesis: A Review of the Literature Regarding Acquisition of a Native-Like Pronunciation in a Second Language. Proceedings from the TRI-TESOL Conference, Bellevue, WA.

Callan, D. E., Tajima, K., Callan, A. M., Kubo, R., Masaka, S., and Akahane-Yamada, R. (2003). Learning-Induced Neural Plasticity Associated with Improved Identification Performance After Training of a Difficult Second-Language Phonetic Contrast. *NeuroImage*, 19(1), 113-124.

Cenoz, J., and Lecumberri, L.G.. The Effect of Training on the Discrimination of English Vowels. (1999). *International Review of Applied Linguistics in Language Teaching,* 37(4), 261-275.

Coulston, R., Oviatt, S., and Darves, C. (2002). Amplitude Convergence in Children's Conversational Speech with Animated Personas. Proceedings from the 7th International Conference on Spoken Language Processing, Denver, CO.

Delattre, P. (1965). *Comparing the Phonetic Features of English, French, German, and Spanish*. Heidelberg: Julius Groos Verlag.

Delvaux, V. and Soquet, A. The Influence of Ambient Speech on Adult Speech Productions Through Unintentional Imitation. *Phonetica* 64, 145-173.

Flege, J.E. and Eefting, W. (1987) Cross-Language Switching in Stop Consonant Perception and Production by Dutch Speakers of English. *Speech Communication,* 6, 185-202.

Flege, J.E., Frieda, E.M., and Nozawa, T. (1997). Amount of Native-Language (L1) Use Affects the Pronunciation of an L2. *Journal of Phonetics,* 25(2), 169-186.

Grosjean, F. (1989). Neurolinguists, Beware! The Bilingual is Not Two Monolinguals in One Body. *Brain Language, 36*(1), 3-15.

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.

Hardcastle, W. J. and Hewlett, N. (Eds). (1999). *Coarticulation: Theory, Data, and Techniques.* Cambridge: Cambridge University Press.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic Characteristics of American English Vowels." *Journal of the Acoustical Society of America* 97(5), 3099-3110.

Hyltenstam, K. and Abrahamsson, N. (2000). Who Can Become Native-Like in a Second Language? All, Some, or None? On the Maturational Constraints Controversy in Second Language Acquisition. *Studia Linguistica,* 15(2), 150-166.

---. (2003). Maturational Constraints in Second Language Acquisition. In C. Doughty and M. Long (Eds.), *Handbook of Second Language Acquisition* (539-588). Oxford: Blackwell.

Jiang, H. The Effect of L2 Phonetic Learning on L1 Vowels. (Doctoral dissertation, Simon Frasier University, 2008).

Lamendella, J. T. (1977). General Principles of Neurofunctional Organization and Their Manifestation in Primary and Nonprimary Language Acquisition. *Language Learning,* 27(1), 155-196.

Lim, V. P. C., Liow, S. J. R., Lincoln, M., Chan, Y. H., and Onslow, M. (2008). Determining Language Dominance in English-Mandarin bilinguals: Development of a Self-Report Classification Tool for Clinical Use." *Applied Psycholinguistics,* 29, 389-412.

Martínez-Celdrán, E., Fernández-Planas, A. M., and Carrera-Sabaté, J. (2003). "Castilian Spanish." *Journal of the International Phonetic Association,* 33, 255-259.

Ohala, J. (1994). "Acoustic Study of Clear Speech: A Test of the Contrastive Hypothesis." Proceedings from the International Symposium on Prosody, Yokohama, Japan.

Oyama, S. (1976). A Sensitive Period for the Acquisition of a Nonnative Phonological System. *Journal of Psycholinguistic Research* 5(3), 261-283.

Pardo, S. (2006). On Phonetic Convergence During Conversational Interaction. *The Journal of the Acoustical Society of America,* 119(4), 2382-2393.

Pikering, M. J. and Garrod, S. (2004). Towards a Mechanistic Psychology of Dialogue. *Behavioral and Brain Sciences,* 27, 169-226.

Piske, T., MacKay, I. R. A., and Flege, J. E.. (2001). Factors Affecting Degree of Foreign Accent in an L2: A Review. *Journal of Phonetics* 29(2), 191-215.

Piske, T., Flege, J. E., MacKay, I. R. A., and Meador, D. (2002). The Production of English Vowels by Fluent Early and Late Italian-English Bilinguals. *Phonetica* 59(1), 49-71.

Sancier, M. L. and Fowler, C.A.. (1997). Gestural Drift in a Bilingual Speaker of Portuguese and English. *Journal of Phonetics,* 25, 421-436.

Sebastián-Gallés, N. and Soto Faraco, S. (1999). Online Processing of Native and Non-Native Phonemic Contrasts in Early Bilinguals. *Cognition* 72(2), 111-123.

Sebastián-Gallés, N., Echeverría, S., and Bosch, L. (2005). The Influence of Initial Exposure on Lexical Representation: Comparing Early and Simultaneous Bilinguals. *Journal of Memory and Language,* 52(2), 240-255.

Street, R. Jr. (1984). Speech Convergence and Speech Evaluation in Fact-Finding Interviews. *Human Communication Research* 11(2), 139-169.

Thompson, I. (2006). "Foreign Accents Revisited: The English Pronunciation of Russian Immigrants." *Language Learning*, 41(2), 177-204.

Whitley, M. S. (2002). Phonemes. In Spanish-English Contrasts: A Course in Spanish Linguistics (pp. 20-41). Washington: Georgetown University.

Williams, L. (1977). The Voicing Contrast in Spanish. *Journal of Phonetics* 5(2), 169-184.

---. (1979). The Modification of Speech Perception and Production in Second-Language Learning." *Perception and Psycholinguistics*, 26(2), 95-104.

Appendix A: Language History Background Survey

I. Identifying information

Name:
Today's Date:
Nationality:
Gender:
Country of Birth:
Date of Birth:
Please list all the places you have lived and the number of years you have lived there:
Please list the following information about your education: Name of school, number of years attended, location of school, primary language of instruction Number of years of formal English instruction: _____ Number of years of formal Spanish instruction: _____
Number of years of exposure to English:
Number of years of exposure to Spanish:

II. Understanding

Please write down a number to show which languages you **UNDERSTAND BEST**. For example, if you understand English best, put the number "**1**" next to the word "English." If you understand Spanish second best, put a number "**2**" next to the word "Spanish." If you cannot understand any of the languages, put a "**0**" next to that language. Also, please report the age at which you started to **UNDERSTAND** each of the languages that you know. For example, you may have started to hear and understand Spanish at home (age = 1 year) but you did not start hear and understand English until kindergarten (age = 5 years). If you cannot remember exactly, make an educated guess.

Language	Ranking	Age of First Exposure
English		
Spanish		
Others (please specify):		

Please **circle** a number on the rating scale below to indicate the proficiency/competency with which you can **CURRENTLY UNDERSTAND** each language. **DO NOT USE** half-points (e.g., 3.5).

1. How proficient are you in understanding English?

"I can only understand a few words."						"I have native proficiency."
1	2	3	4	5	6	7

2. How proficient are you in understanding Spanish?

"I can only understand a few words."						"I have native proficiency."
1	2	3	4	5	6	7

3. How proficient are you in understanding other languages (specify _____)?

"I can only understand a few words."						"I have native proficiency."
1	2	3	4	5	6	7

4. How proficient are you in understanding other languages (specify _____)?

"I can only understand a few words."						"I have native proficiency."
1	2	3	4	5	6	7

III. Speaking

Please write down a number to show which languages you **SPEAK BEST**. For example, if you understand English best, put the number "**1**" next to the word "English." If you understand Spanish second best, put a number "**2**" next to the word "Spanish." If you cannot understand any of the languages, put a "**0**" next to that language. Also, please report the age at which you started to **SPEAK** each of the languages that you know. For example, you may have started to hear and understand Spanish at home (age = 1 year) but you did not start hear and understand English until kindergarten (age = 5 years). If you cannot remember exactly, make an educated guess.

Language	Ranking	Age of First Exposure
English		
Spanish		
Others (please specify):		

Please **circle** a number on the rating scale below to indicate the proficiency/competency with which you can **CURRENTLY SPEAK** each language. **DO NOT USE** half-points (e.g., 3.5).

1. How proficient are you in speaking English?

"I can only speak a few words."						"I have native proficiency."
1	2	3	4	5	6	7

2. How proficient are you in speaking Spanish?

"I can only speak a few words."						"I have native proficiency."
1	2	3	4	5	6	7

3. How proficient are you in speaking other languages (specify _____)?

"I can only speak a few words."						"I have native proficiency."
1	2	3	4	5	6	7

4. How proficient are you in speaking other languages (specify _____)?

"I can only speak a few words."						"I have native proficiency."
1	2	3	4	5	6	7

Appendix B: Stimuli Sentences and Images

Sentences for Spanish-accented English session

There is one telephone on the shelf.
There are two telephones on the shelf.
There are three telephones on the shelf.

There is one telephone on the table.
There are two telephones on the table.
There are three telephones on the table.

There is one telephone on the boat.
There are two telephones on the boat.
There are three telephones on the boat.

There is one rose on the shelf.
There are two roses on the shelf.
There are three roses on the shelf.

There is one rose on the table.
There are two roses on the table.
There are three roses on the table.

There is one rose on the boat.
There are two roses on the boat.
There are three roses on the boat.

There is one sheet of paper on the shelf.
There are two sheets of paper on the shelf.
There are three sheets of paper on the shelf.

There is one sheet of paper on the table.
There are two sheets of paper on the table.
There are three sheets of paper on the table.

There is one sheet of paper on the boat.
There are two sheets of paper on the boat.
There are three sheets of paper on the boat.

There is one shell on the shelf.
There are two shells on the shelf.
There are three shells on the shelf.

There is one shell on the table.
There are two shells on the table.
There are three shells on the table.

There is one shell on the boat.
There are two shells on the boat.
There are three shells on the boat.

There is one photo on the shelf.
There are two photos on the shelf.
There are three photos on the shelf.

There is one photo on the table.
There are two photos on the table.
There are three photos on the table.

There is one photo on the boat.
There are two photos on the boat.
There are three photos on the boat.

There is one weight on the shelf.
There are two weights on the shelf.
There are three weights on the shelf.

There is one weight on the table.
There are two weights on the table.
There are three weights on the table.

There is one weight on the boat.
There are two weights on the boat.
There are three weights on the boat.

Sentences for English-accented Spanish session

Hay uno teléfono en la banqueta.
Hay dos teléfonos en la banqueta.
Hay tres teléfonos en la banqueta.

Hay uno teléfono en la mesa.
Hay dos teléfonos en la mesa.
Hay tres teléfonos en la mesa.

Hay uno teléfono en el bote.
Hay dos teléfonos en el bote.
Hay tres teléfonos en el bote.

Hay uno rosa en la banqueta.
Hay dos rosas en la banqueta.
Hay tres rosas en la banqueta.

Hay uno rosa en la mesa.
Hay dos rosas en la mesa.
Hay tres rosas en la mesa.

Hay uno rosa en el bote.
Hay dos rosas en el bote.
Hay tres rosas en el bote.

Hay uno boleta en la banqueta.
Hay dos boletas en la banqueta.
Hay tres boletas en la banqueta.

Hay uno boleta en la mesa.
Hay dos boletas en la mesa.
Hay tres boletas en la mesa.

Hay uno boleta en el bote.
Hay dos boletas en el bote.
Hay tres boletas en el bote.

Hay uno bandera en la banqueta.
Hay dos banderas en la banqueta.
Hay tres banderas en la banqueta.

Hay uno bandera en la mesa.
Hay dos banderas en la mesa.
Hay tres banderas en la mesa.

Hay uno bandera en el bote.
Hay dos banderas en el bote.
Hay tres banderas en el bote.

Hay uno foto en la banqueta.
Hay dos fotos en la banqueta.
Hay tres fotos en la banqueta.

Hay uno foto en la mesa.
Hay dos fotos en la mesa.
Hay tres fotos en la mesa.

Hay uno foto en el bote.
Hay dos fotos en el bote.
Hay tres fotos en el bote.

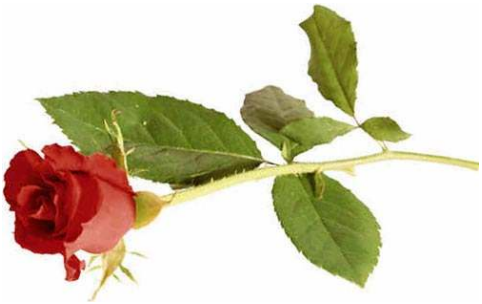
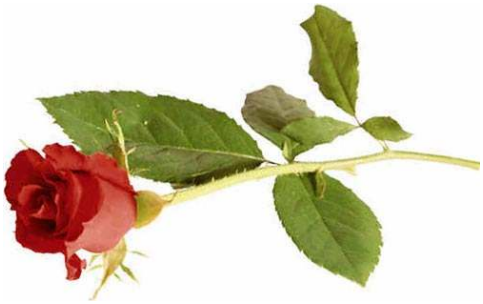
Hay uno pesa en la banqueta.
Hay dos pesas en la banqueta.
Hay tres pesas en la banqueta.

Hay uno pesa en la mesa.
Hay dos pesas en la mesa.
Hay tres pesas en la mesa.

Hay uno pesa en el bote.
Hay dos pesas en el bote.
Hay tres pesas en el bote.

Stimuli Images and Words

English	**Spanish**
T**e**lephone	Tel**é**fono
	
Sh**e**ll	Band**e**ra
	
Location: sh**e**lf	Location: banqu**e**ta
	

R**o**se	R**o**sa
	
Ph**o**to	F**o**to
	
Location: b**o**at	Location: b**o**te
	

W**ei**ght	P**e**sa
	
Sheet of p**a**per	Bol**e**ta
	
<u>Location</u>: t**a**ble	<u>Location</u>: m**e**sa
	

Appendix C: Post-Test Phonetic Assessment

We are interested in your personal perception of the way you speak, your accent in English and Spanish, and the effect that you perceive the speech of others to have on the way that you speak.

Please circle the number that best corresponds to your impression of how you speak.

1. How do you judge your accent to be in English?

1 2 3 4 5 6 7
(1 – "I speak English with a very heavy foreign accent." 7 – "I speak English like a native speaker of English.")

Comments:

_____
_____
_____

2. How do you judge your accent to be in Spanish?

1 2 3 4 5 6 7
(1 – "I speak Spanish with a very heavy foreign accent." 7 – "I speak Spanish like a native speaker of Spanish.")

Comments:

_____
_____
_____

3. How much do you imitate the speech sounds of others, intentionally or otherwise?

1 2 3 4 5 6 7
(1 – "Never." 7 – "Always.")

Comments:

_____
_____
_____

Are you aware when this happens?: ☐ Yes, I notice my accent changing based on the speech I hear.
☐ No; I only recognize this when I am told that I am doing it.

Appendix D – Correlations

As described in Chapter 3, correlations were run to compare responses to question 1 and 2 to production during test 1 and test 2 sections of the experiment of duration and formants. The significant correlations are as follows:

Spanish-accented English Session		
Accent Self-perception (Question 1)		
Correlated with:	Correlation	Significance (p =)
e/ei test 2 formant 3 change	- .491	.038
Degree of Convergence (Question 3)		
Correlated with:	Correlation	Significance (p =)
o/ou test 2 formant 2 change	- .491	.038
o/ou test 2 formant 1 change	-.770	< .001
o/ou test 1 formant 1 change	- .729	.001
e/ei test 2 formant 3 change	- .593	.009
e/ei test 2 formant 2 change	- .659	.003
English-accented Spanish Session		
Accent Self-perception (Question 2)		
No significant correlations		
Degree of Convergence (Question 3)		
Correlated with:	Correlation	Significance (p =)
o/ou test 1 duration	- .613	.026
o/ou test 2 duration	- .578	.039
e/ei test 1 duration	- .592	.033
e/ɛ test 2 formant 1 change	.594	.032

From the chart above, it is evident that there are some correlations between production during the test phases and the degree to which participants reported converging to speech in conversation.

For the Spanish-accented English session, the correlations indicate that as participants' self-reported tendency to converge (as measured by question 3) increased, their formant change measures (the amount that their vowels diphthongized) decreased, which aligns with predictions that participants who converge more would show increased monophthongization, in line with the production of the accented speech they heard.

In the English-accented Spanish session, the negative correlation between question 3 responses and durations indicates that as convergence increased, durations decreased – a result that goes against predictions, as it means that participants who self-rated as more likely to converge actually showed greater production of shorter durations for the vowels specified, even though longer durations would be predicted in the English-accented Spanish section as a result of the longer vowels produced by the American English-accented speaker. Only "e/ɛ test 2 formant 1 change" was positively correlated with convergence, meaning that as convergence tendencies increased, formant 1 movement in this vowel also increased, signaling an increase in diphthongization, which aligns with predictions of converge. However, given the problem with formant 1 analysis discussed in Chapter 3, even this result is questionable.

For accent self-perception (questions 1 and 2), there was only one positive correlation: in the Spanish-accented English session, e/ei test 2 formant 3 change was positively correlated with the degree to which participants felt their Spanish accent was native-like. This means that as participants rated themselves as more native-like in their English production, their formant 3 production for e/ei during the test 2 phase was increasingly diphthongized. This result conforms with expectations that a native English-like accent would lead to increased diphthongization in vowels.

Discussion

I ran correlations between post-test assessment questions and performance during the experiment to determine whether there was a consistent pattern of results demonstrating that self-perception of speech is related to unconscious convergence during the experimental sessions. However, the pattern of results that emerged was inconsistent, incomplete, and contradictory at times. The more consistent set of correlations is between formant changes in the Spanish-accented English condition and participants' self-rating of degree of convergence (question 3). These correlations demonstrate that participants who rated themselves are more consistent convergers were those who showed smaller formant change values (i.e. more monophthongization) in the test 1 and test 2 phases of the Spanish-accented English session. However, the measures that did produce positive correlations are scattered throughout the formants and the different vowels, and do not produce a clear picture of what, exactly, high-converging participants were converging on with respect to the accented speech they heard.

The English-accented Spanish data is even more troubling, as the correlations go in the opposite direction of what would be predicted. For some duration measures, there was a negative correlation between self-rating of convergence and duration, meaning that those who considered themselves more prone to convergence had shorter durations at time 1 and time 2. However, the English-accented speaker produced longer durations, and thus the direction of correlation in these measures seems to be backwards. The only positive correlation for data during the English-accented Spanish session was between degree of convergence and e/ɛ test 2 formant 1 change.

Because the correlations in this session do not tell a cohesive story, and because the significantly correlations are scattered across times and vowels in a seemingly random distribution, it may be the case that the question asking participants to rate their degree of

convergence was not a particularly useful measure in predicting how much they will converge in the experimental setting. In other words, an individual's *impression* of how much he or she converges in a conversational context may not have much to do with how much he or she *actually* converges to accented speech that he or she hears in an experimental setting. In other words, individuals may not be good at tracking their tendencies to phonetically convergence (at least in an experimental setting).

The one positive correlation between accent self-perception ("How native do you believe your accent is in Spanish/English?") and experimental production was in the Spanish-accented English condition. There was a positive correlation between accent self-perception and diphthongization in formant 3 in the vowel e/ei during the test 2 phase. This means that participants who rated themselves as having a native-like English accent also produced vowels with increased diphthongization in the second formant of the vowel e/ei during the second test phase. This correlation makes sense, as speakers who have diphthongized vowels will sound more like American English speakers. So for this particular measure, the self-perception rating seems valid. However, the failure to achieve any other significant correlations suggests that the measure of self-perception of accent was not a strong predictor of performance during the experiment.

Future studies that attempt to find correlations between performance and awareness of phonetic convergence may need to rely on more robust measures of awareness. In addition to asking participants their own opinions of the nativeness of their accents, it would be useful to employ outside listeners to give an independent rating of paritcipants' accents. Because individuals may have a skewed or inaccurate perception of the way they speak as compared to how others view their speech, this extra measure would help to discover any inconsistencies in

self-perception and perhaps provide a measure that correlates more strongly with performance during the experimental task.