View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Scholar Commons | University of South Florida Research

University of South Florida
Scholar Commons

Graduate Theses and Dissertations

Graduate School

2005

The effects first language use phonological difficulty perception foreign accented speech [sic]

Astrid Zerla Doty
University of South Florida

Follow this and additional works at: http://scholarcommons.usf.edu/etd

Part of the American Studies Commons

Scholar Commons Citation

Doty, Astrid Zerla, "The effects first language use phonological difficulty perception foreign accented speech [sic]" (2005). *Graduate Theses and Dissertations.*
http://scholarcommons.usf.edu/etd/2860

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

The Effects First Language Use Phonological Difficulty

Perception Foreign Accented Speech

by

Astrid Zerla Doty

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts
Department of Psychology
College of Arts and Sciences
University of South Florida

Co-Major Professor: Ruth Huntley Bahr, Ph.D.
Co-Major Professor: Judith Becker Bryant, Ph.D.
Catherine Rogers, Ph.D.
Doug Nelson, Ph.D.

Date of Approval:
February 4, 2005

Keywords: second language acquisition, bilingual, phonology

© Copyright 2005, Astrid Zerla Doty

Dedication

To my family who believed in me so much, they left no room for my own doubts. All that I am, and all that I hope to be, I owe to them.

"Somewhere, something incredible is waiting to happen."

Carl Sagan

Acknowledgements

I wish to express my gratitude for the scholarly advise and guidance of Dr. Ruth Bahr. Her encouragement and support helped through many times.

I also wish to extend special thanks to Dr. Judy Bryant, Dr. Cathy Rogers, and Dr. Doug Nelson who were always available for assistance and assurance.

Thanks also to my family and friends who believed in and helped me in so many invaluable ways.

Table of Contents

List of Tables iii

List of Figures iv

Abstract v

Introduction 1
- Critical Period Hypothesis 3
- Equivalence Classification Hypothesis and the Single System Hypothesis 6
- ECH Rests Upon Single System Assumptions 7
- How Attitude Influences L2 Learning and Foreign Accent 9
- Social Factors That Influence L2 Learning 13
- Speaker Variables 13
  - Age of Learning 13
  - Amount of Native Language Use 17
  - Experience 18
- Speaker and Listener Variables 19
- Characteristics of the Stimuli 21
  - Temporal and Acoustic Properties of the Stimuli 22
  - Speaking Rate 22
  - Length of Stimuli 23
  - Rating Scales 23
  - Linguistic Variables 25
  - Phonological Characteristics of the L1 and the L2 25
- Purpose of the Proposed Study 28

Method 31
- Design 31
- Participants 31
  - Speakers 31
  - Listeners 32
  - Materials 34
    - Demographic questionnaire for the speakers 34
    - Demographic questionnaire for the listeners 34
    - Practice sentences 36
    - Stimulus Recording 37
    - Recording of the stimulus sentences 38
    - Stimulus preparation 40
- Procedure 41

Data Reduction 43

Results 43
Intra-rater Reliability 45
Degree of Perceived Accent 45
Analysis of Individual Speakers Ratings 50
Analysis of Speakers' Production of Stimuli Targets 53
Summary 62

Discussion 64
Differences Between Listener Groups 65
Characteristics of the Stimuli 71
Future Directions 74
Conclusions 76

References 78

Appendices 83
Appendix A: Study Advertisement 84
Appendix B: Practice Sentences 85
Appendix C: Stimulus Sentences 86
Appendix D: Instructions to Speakers 88
Appendix E: Instructions to Listeners 89

List of Tables

Table 1	Speaker Characteristics	32
Table 2	Listener Characteristics	34
Table 3	Distribution of Targets for Less and More Difficult Sentences	55

List of Figures

Figure 1 Mean Accentedness Ratings by Listener and Speaker 47

Figure 2 Mean Ratings of Accent by Both Speaker and Listener Groups 49

Figure 3 Mean Accentedness Ratings of Each Speaker 51

Figure 4 Mean Ratings by Listener and Heavily and Moderately Accented Speaker Group 52

Figure 5 Mean Accent Ratings for Individual Speakers for Less and More Difficult Sentences 53

Figure 6 Mean Percent Target Errors by Individual Speakers and Sentence Type 54

Figure 7 Mean Percent Target Errors for Less and More Difficult Sentences 56

Figure 8 Mean Percent Errors by Individual Speakers for Less and More Difficult Sentences 61

First Language Use and Phonological Difficulty on the
Perception of Foreign Accented Speech

Astrid Zerla Doty

ABSTRACT

Listener perception of accentedness has been shown to be influenced by experience with L2 (measured by length of residence in US). However, frequency of L1 use and degree of phonological difficulty (defined by the number of non-native phonetic features targeted) may provide more insight into the role of experience in the perception of accentedness.

Three groups of listeners (monolingual English and Spanish [L1] speakers divided into two groups of high and low use of English [L2]) rated the accentedness of bilingual speakers who spoke with varying degrees of accentedness. The speakers read sentences adapted from Magan (1998) to include phonological aspects likely to be difficult for native Spanish speakers.

Listeners performed similarly in rating speakers' degree of accent. Amount of daily L1 use only influenced the ratings of the slightly accented group; the high-use bilingual group rated these speakers as more accented than the native English group, regardless of level of phonological difficulty. These results suggest that the high-use groups' lack of L2 experience made them less perceptually sensitive to certain phonetic features of English. Because speakers did not make the predicted target errors, the listener groups may have based their ratings on features not targeted in this investigation

Chapter 1

Introduction

Evidence suggests that in the domain of phonology, the younger a person learns a second language, the more likely he or she will be able to pass as a native speaker in that language, but the reasons why this happens have been debated for some time. The theories used to explain the effects of age on the learning of a second language have evolved from those that are more neurologically based to those that build upon these neurological models with the addition of sociological considerations. Two major theoretical approaches dominate the literature as explanations of how age-related factors influence the learning of a second language: the critical period hypothesis (CPH) (Lenneberg, 1967) and equivalence classification hypothesis (ECH) (Bohn & Flege, 1990; Flege & Eefting, 1987). In this paper, the ECH will be considered within the framework of the single-system hypothesis.

The earlier of the two theories, the CPH, states that there are neurological and maturational constraints that influence the learning of a second language. However, Flege contends that the ability to learn a second language remains intact across the lifespan. The ECH deals with the interaction of the phonological systems of the first and second languages in predicting areas of difficulty for the second language learner and recognizes sociocultural and sociolingistic factors as influential in the production of foreign accent.

These are areas the critical period hypothesis ignores, yet have been shown to be highly related to foreign accent, regardless of the age of learning.

Numerous investigations have documented patterns of second language acquisition in children and adults. Many have credited the age differences in L2 acquisition to changes in the physiology of the brain alone, without probing more deeply as to how these changes specifically relate to the learning of phonology in the L1 and the L2. Certainly, the age at which one learns a second language is much more complex than previously thought.

In the initial section of this paper, the CPH is explained and its limitations described. Following this, the ECH and the single-system hypothesis are considered because they represent an evolution of the critical period hypothesis and a paradigm shift in the understanding of second language phonological learning. In the present study, the degree of perceived foreign accent was evaluated in speakers with varying degrees of accentedness by listeners who differed in their amount of L1 use. The speakers read sentences that varied in phonological difficulty, which was manipulated by creating sentences that include varying numbers of targets that are deemed difficult for Spanish speakers who speak English as a second language to produce. Those evaluating the sentences were also bilingual and differed in their amount of daily L1 use.

Critical Period Hypothesis

The older and more traditional view of second language learning stems from Eric Lenneberg's (1967) argument that a critical period exists for the acquisition of a first language (L1). Lenneberg proposed that this period extends from about two years of age through the end of puberty, which he marked at age 14. One criticism of this hypothesis

is that Lenneberg only considered first language acquisition; there is no definitive or widely accepted theory regarding a critical period for the acquisition of a second language. Further, if one accepts the notion that a critical period exists for second language learning, the question remains as to what the boundaries are for this period.

Lenneberg's critical period hypothesis was formed at a time when little evidence was available to test it directly. That is, there existed no credible reports of normal children who had been deprived of exposure to a first language. Therefore, he based his hypothesis on indirect evidence, such as differences in recovery from aphasia for children versus adults, and differences in the progress of language acquisition before and after puberty in children who were mentally retarded. He claimed that neurological underpinnings were responsible for the maturational changes observed in language learning abilities. Lenneberg suggested that after puberty, the brain loses the plasticity and organizational capacities necessary for acquiring language. The implication is that any language acquisition that takes place after puberty will be qualitatively different from that involved in first language acquisition. By extension, any language learning that occurs after the age of puberty will be more laborious and less successful (Lenneberg, 1967).

There are degrees to which researchers in second language learning have subscribed to the CPH, based mainly on the extent to which the theory accounts for exceptional cases of adult second language learning. One interpretation of the CPH is the strong version (Neufeld, 1979). Briefly stated, the assumptions of this position are that there are biological constraints upon second language learning in adults, that these constraints are inevitable and irreversible, and that no one beyond puberty can hope to

lose his or her foreign accent in the second language. The soft or weak version, to which Lenneberg ascribed, states that *most* adults will be incapable of native-like speech in the second language. In addition to the strong and weak versions of the CPH, there exist further variations.

The term "sensitive period", which is similar to the weak version, refers to the notion that the age limitation on second language acquisition is not absolute in the sense proposed by the critical period hypothesis (Patkowski, 1980). Rather, the approach suggests that is possible to acquire a second language after the sensitive period, but it would not be possible to attain native-like proficiency. Patkowski (1980) suggested that the term "critical period" be reserved for cases of first language acquisition, while the term "sensitive period" be used in the case of second language acquisition, because the limitation is on the ability to acquire complete native-like proficiency in the foreign language.

For those in the critical period camp, there is a difference of opinion as to the range on maturational constraints on second language learning. Some argue that the range of age-related constraints is limited only to phonology, while others contend it extends into other domains of language, such as syntax, morphology, and semantics. Adults may have a better ability to think about language and use for their learning of an L2 some of the same skills they acquired in learning and mastering their L1. Yet, for reasons not fully understood, adults apparently initially acquire a second language faster than young children, yet the child learners eventually achieve more native-like mastery of the L2 that adults rarely experience (Long, 1990; Snow & Hoefnagel-Hohle, 1978). Additionally, the initial advantages that adults experience during L2 learning seem not to

involve the domain of phonology, but are restricted to other domains of language such as syntax, morphology, and the lexicon.

Equivalence Classification Hypothesis and the Single System Hypothesis

The CPH is limited in that it does not fully consider how maturational constraints interact with sociocultural variables. A more useful explanation is termed the single system hypothesis, (Flege, Freida, & Nozawa, 1997) which asserts that bilinguals have a single phonological system in which both their languages reside and that they cannot fully isolate either phonetic system (Guion, Flege & Loftin, 2000). It further predicts the loss or attenuation of L1 through disuse. In other words, the less L1 there is, the smaller will be its influence on the L2 (Flege et al., 1997). According to Grosjean (1992), the L1 phonetic system influences that of the L2, and the nature of this influence depends on several variables, including the amount and type of use of each language. Generally, the more individuals speak their native language, the stronger will be their accent in their second language (Flege et al., 1997; Guion et al., 2000). Furthermore, this relationship seems to be asymmetrical; although L1 use has an effect on accent in the L2, the L2 has little effect on L1 production. The single system hypothesis also states that the loss of L1 may reduce the degree of perceived foreign accent in an L2.

The single system hypothesis, which makes predictions based on the amount of use of the L1 and L2, is enhanced when one considers the ECH. This hypothesis distinguishes between identical, similar, and new sounds in a cross-language context (Bohn & Flege, 1990; Flege & Eefting, 1987). First, consider the perceptual assimilation model (PAM), which asserts that during L1 speech acquisition, non-native segments tend to be perceived according to their similarities to and/or differences from the closest native

speech segment (Best, 1995). According to this model, listeners will perceptually assimilate non-native phones to native categories. The equivalence classification hypothesis, on the other hand, deals more specifically with the perceptual assimilation of second language phones to native categories. The predictions of the hypothesis are that identical sounds in two languages (e.g., the German and English /b/) are unlikely to cause a problem, but similar sounds, like the English and German /u/, might offer persistent although subtle problems for the second language learner in acquisition. Similar sounds, therefore, should be most difficult because they will probably be substituted by the sound from the first language, even after extended L2 exposure. Sounds that are completely new, in the sense that they are not equivalent or even similar to sounds from the individual's L1, will be established into a new category as a result of phonetic learning that is not hampered by equivalency classification. As the amount of experience with the L2 increases, individuals will produce second language vowels more like natives (Flege et al., 1997). However, production varies as a function of the relationship between the native and second language phonology. Therefore, the single system hypothesis seems to work in conjunction with the ECH in that both consider experience with the L2 and the interactions of the L1 and L2 phonetic systems.

ECH Rests Upon Single System Assumptions

Opinion among researchers as to the existence of maturational constraints in second language learning is sharply divided, with both sides offering supporting evidence. Several researchers have shown what they considered to be such evidence of maturation constraints operating (Birdsong, 1992; Johnson & Newport, 1989, 1991; Patkowski, 1980; Tahta, Wood, & Loewenthal, 1981). Others have claimed that their

findings suggest an advantage for older learners and rejected the CPH altogether, even with respect to pronunciation or phonology (Hill, 1970; Neufeld, 1979; Snow & Hoefnagel-Hohle, 1978). Generally, the literature thus far has supported three generalizations: adults proceed through early stages of morphological and syntactic development more quickly than children do, older children acquire these domains more quickly than younger children, and child learners outperform post-pubescent learners in the long run (Long, 1990).

To test the CPH with specific attention to rate of acquisition, Snow and Hoefnagel-Hohle (1978) conducted a longitudinal study of the natural acquisition of Dutch by English speakers of different ages. The authors tested two groups English speakers: monolingual English speakers who were just starting to learn Dutch and English speakers who had been living in the Netherlands and speaking Dutch for at least 18 months. The beginning learners were tested three times at four to five month intervals. The advanced learners were tested only once. The beginning learners wree distributed into the following age groups: three to five year-olds, six-seven year-olds, eight-ten year-olds, 12-15 year-olds, and adults. The advanced learners were distributed into the following age groups: six-seven year-olds, eight-ten year-olds, 12-15 year-olds, and adults. Participants were assessed in the areas of imitative and spontaneous pronunciation, auditory discrimination, morphology, sentence repetition, sentence translation, sentence judgment, vocabulary, story comprehension, and storytelling. The results of this study point to faster initial learning in the older subjects relative to the younger ones, but not differences in ultimate attainment. Interestingly, there were also differences noted within the group of native-speakers on morphology and auditory

discrimination tasks. The authors contended that differences noted within the native-speaker group are important to the assessment of the CPH. If native-speakers demonstrate a range of skills in their first language, then it seems logical that, by extension, post-adolescent second language learners will not achieve equal skills in their second language.

The cause for the range of skills demonstrated by second-language learners who began L2 acquisition at the same age is not addressed by the CPH. One of the variables believed to contribute to a speaker's degree of foreign accent, regardless of age of acquisition, is attitude (Anderson & Koehler, 1988; August & Hakuta, 1998; Bresnahan, Ohashi, Nebashi, Liu, & Shearman, 2002; Cummins, 2000; Hill, 1970; Zecker, 2004). Attitude towards one's second language may indirectly affect one's foreign accent in that it determines, to a large extent, the amount of daily L1 use, a variable found to contribute significantly to accent under the EC and the willingness to lose the accent.

How Attitude Influences L2 Learning and Foreign Accent

Disputes over the CPH stem from researchers who contend that age constraints are not only due to neurological changes but may reflect social factors. They insist that the disparity between child and adult performance can better be explained by social and psychological factors that are independent of psycholinguistic abilities but dependent on cultural tradition (Anderson & Koehler, 1988; August & Hakuta, 1998; Bresnahan, Ohashi, Nebashi, Liu, & Shearman, 2002; Cummins, 2000; Hill, 1970; Zecker, 2004). These factors include status of the first and second languages, motivation or the extent to which one needs to learn the second language in order to function in the target or second culture, and the cognitive demands of learning a second language.

There is some reason to doubt that the advantages children seem to have in attaining mastery of second languages are uniform across cultures. Hill (1970) explored the influence of the social and cultural aspects of language, as well as attitudes surrounding second language use. For example, she pointed out that adults have opinions about the negative and positive qualities of a second language. These ideas are certainly not inborn, but are the result of their cultural traditions (August & Hakuta, 1998). For example, Hill (1970) cited studies that examined the role and nature of second language learning in American-Indian and Australian New Guinea cultures. In these studies, adults acquired new languages because of the roles multilinguilism played in political activities. Thus, a motivational factor has been identified as contributing to language learning success. Hill noted that most of these early ethnographic studies did not examine the question of language proficiency or whether adults master foreign languages as well as children in communities where there is intense social and political pressure for adults to learn another language. Likewise, August and Hakuta (1998) found that the extreme importance of learning English in order to succeed in American society overrode immigrants' negative attitudes towards English. Additionally, motivation is an important factor in Americans' apparent lack of bilingualism. There seems little reason to learn a second language when English is considered by many Americans to be a "world language" because anywhere they are in the world, someone will speak English.

Another attitude potentially affecting the learning of a second language is that multilingualism is bad for children (August & Hakuta, 1998; Hill, 1970; Zecker, 2004). The contention is that children exposed to more than one language will not perform as well on intelligence tests compared to monolingual children. However, Zecker found that

English- and Spanish-dominant children placed in a two-way immersion classroom (one in which instruction was in both Spanish and English) actually performed considerably better on English literacy achievement measures than did English speakers in regular monolingual English classrooms.

Unfortunately, children who would naturally be expected to be bilingual, such as American-Indian and Mexican- or Cuban-Americans, may find their bilingualism discouraged in the school setting, partly because the cultures with which these languages are associated are considered by many teachers to be lower class (August & Hakuta, 1998; Hill, 1970). Therefore, these children may experience the loss of their first language, which subsequently would affect their degree of foreign accent. Conversely, adult bilinguals are often considered to be exceptionally intelligent, but again, only if their language is associated with a high-status culture. Thus, these individuals may be motivated to maintain their foreign accent. The high-status given to some languages and not others also may be responsible for reinforcing the idea that adults can never lose their foreign accent. This distinction between the linguistic majority and minority was discussed by August and Hakuta (1998), who described the effects of societal variables, such as prestige and status of the languages, involved in bilingualism. Their conclusion was that immigrants whose language was not valued in the United States experienced erosion of their first language, including its phonology, which influenced their degree of foreign accent.

Although Hill (1970) made her observations more than 30 years ago, more recent studies lend credence to her assertions. Bresnahan et al. (2002) evaluated attitudinal and affective responses toward accented English as a function of speakers' identity and

intelligibility. The authors also sought to determine whether participants' level of ethnic identity had any relationship to their attitude towards accented speech. Native English speakers from various ethnic backgrounds listened to recorded messages in one of six conditions: intelligible foreign friend, intelligible foreign teaching assistant, unintelligible foreign friend, unintelligible foreign teaching assistant, intelligible American friend, and intelligible American teaching assistant. After listening to the recordings, participants then completed a 101-item questionnaire which assessed attitude, affective response, and ethnic identity. American English was the preferred accent, followed by intelligible foreign accent, with unintelligible foreign accent the least preferred. Role also influenced participants' attitude: friends evoked more positive responses than teaching assistants in all conditions. These responses may have been a result of confounds in stimuli used. The script for the friend was a narrative describing a trip to visit a roommate's family, whereas the teaching assistant script was a lecture on human communication. It is possible that the friend's script was inherently more desirable to listen to. Interestingly, participants with strong ethnic identity deemed American English, which reflected their ethnicity, to be more pleasing and have higher status than the unintelligible foreign accent, whereas those with weak ethnic identity found unintelligible foreign accent more pleasing and attributed higher status to it than to American English. The authors explained that people with strong ethnic identity may be more attached to their ethnic group and, therefore, will be more likely to recognize a foreign accent as representing someone in an out-group and have more negative attitudes towards those with foreign accents.

The attitudes found in Bresnahan et al. (2002) echo those described by Hill (1970) and may contribute to Americans' lack of tolerance for *some* foreign accents. It would have been helpful if Bresnahan et al. had listed the first languages of the speakers in their study, because this would have addressed Hill's assertions that some accents are more prestigious than others. These limitations call for the need to examine foreign accent in terms of the ECH, which takes these social factors into consideration.

Social factors that influence L2 learning

Investigators have also evaluated the CPH using other measures, such as age of L2 learning (AOL), age of arrival (AOA) in the target country, foreign language experience, amount of native language use, length of residence, familiarity, and speaking rate (Bohn & Flege, 1990; Flege, 1988; Flege, Bohn, & Jang, 1997; Flege, Frieda, & Nozawa, 1997; Flege, MacKay, & Meador, 1999; Flege, Yeni-Komshian & Liu, 1999b; Gass & Varonis, 1984; Guion, Flege, & Loftin, 2000; Matsura, Chiba, & Fujieda, 1999; Munro & Derwing, 1998). All of these variables have been found to have measurable effects on foreign accent. For example, the later individuals arrive in the country of the second language, the stronger their accent as judged by native listeners (Flege et al., 1999a; Flege et al., 1999b).

Speaker Variables

Age of learning. An individual's degree of foreign accent depends on the age at which second language learning begins (Flege, 1988; Flege & Fletcher, 1992; Munro, Flege, & MacKay, 1996; Tahta et al., 1981). For example, Tahta et al. (1981) looked at predictors of transfer of accent from the first language (L1) to the second language (L2) in a group of people whose acquisition of English as an L2 had begun at ages ranging

from 6 to 15 years. The L1's of the speakers were: Arabic, Armenian, Cantonese, Konkan, Czech, Dutch, French, German, Gujarati, Greek, Hindi, Japanese, Persian, Polish, Portuguese, Spanish, Swahili, Swedish, Urdu, and Serbo-Croatian. Speakers read a paragraph of English into a tape-recorder that was later listened to by three independent native-English speaking judges. The recordings were rated for degree of accent and assigned a score of either 0 (no foreign accent), 1 (slight foreign accent), or 2 (marked foreign accent). There was a significant effect of age of acquisition of English as an L2 on whether there was a foreign accent. Subjects who had learned the L2 by age 6-7 showed no foreign accent. The authors contended that those who commenced leaning the L2 during their 7th to 9th years tended to show very slight, if any, foreign accent. However, the chances of speaking the L2 accent free dropped significantly for those who commenced learning the L2 between the ages of nine and 11 years.

Although Tahta et al.'s study (1981) supports the CPH, some methodological issues are worth noting. First, the stimuli consisted of a paragraph from an airline leaflet, with no concern for the contributions and interactions of the two language systems under consideration. The interaction between the L1 and the L2 may have been especially important for this study, which included subjects with over 20 different native languages. Certain speech sounds that exist in the second language may not exist in the first language, so it would be expected that these sounds might have contributed to the degree of accent. Therefore, if stimuli included speech sounds that were easier for some participants than others, then these speakers may have been judged as less accented compared to that of speakers whose native language phonology differed significantly from language of the stimuli. The degree of accentedness would therefore have been

partially due to the interaction of the languages, rather than the age of L2 learning alone. The authors only controlled for age at which L2 learning began, but did not control for amount of language use, age of arrival in the target culture, or other variables found to be significant influences on foreign accent. Thus, conclusions regarding transfer of accent from the L1 to the L2 must be drawn with caution. Additionally, the rating scale used consisted of only a three-point scale, thus reducing potential variability and categorizing subjects together who might actually have very different degrees of accented L2.

Similar results involving the age of learning (AOL) have been found in studies with better methodologies and more controlled participant groups (Flege, 1988; Flege & Fletcher, 1992; Munro et al., 1996). Flege and Fletcher (1992) found that native Spanish-speaking participants who commenced learning English at the age of five years or earlier could produce English without a detectable accent as measured by the ratings given by native English speakers. In contrast, Chinese subjects with an age of learning (AOL) of 7.6 years spoke with a measurable accent (Flege, 1988). Although the authors of these studies observed that the adults' pronunciation of the L2 improved over time, they concluded that a sensitive period for speech learning is reached long before the age of 12 years. In other words, the age at which a foreign accent first becomes perceptible occurs long before puberty. They further argued that L2 learners of all ages remain remarkably able to establish new phonetic categories for L2 sounds that do not exist in their L1. However, the ability to establish categories for sounds that are similar between the L1 and L2 seems to decrease after the age of five to six years. Instead, learners tend to perceive and produce these similar L2 sounds as the corresponding L1 sounds because they ignore the acoustic differences that distinguish the pairs.

Abundant evidence exists that individuals cannot achieve a native-like accent in a second language unless they are exposed to it at an early age, but some researchers have argued that adult learners can do so and show evidence. Neufeld (1979) sought to determine the extent to which adults could reproduce prosodic and articulatory features of a new language and ultimately demonstrated that high levels of accuracy in pronunciation and intonation are achievable by adult second language learners.

In Neufeld's (1979) first study, after receiving 18 hours of intensive training in Japanese and Chinese phonology, 20 adult native speakers of English practiced five times and then recorded ten phrases of four to eight syllables in length in both languages. The tapes were rated for degree of foreign accent by native speakers of each language. Of the 20 participants, three earned a native speaker rating in one language and one did so in both languages.

Believing that this evidence was not enough to refute the strong version of the CPH, Neufeld investigated individuals who learned their second language as adults and could pass as native speakers of that language. For this second study, 150 French words were prepared that included phonemes and phoneme clusters that were known to be especially difficult for English language learners. Three native French speakers and seven nonnative French speakers recorded the words for judgement by native French speakers. Some nonnative speakers were good enough to be classified as native by some of the judges. Neufeld's findings led him to claim that accent-free second language performance is possible in adulthood and, therefore, there is no critical period for second language acquisition.

Perhaps Neufeld (1979) overstated his case. According to Long (1990), Neufeld's studies suffer from some important limitations and possible methodological flaws. For example, in the study on French (1990), Long raised the question of population validity or generalizability. The nonnative speakers, drawn from a bilingual environment, considered themselves highly proficient and survived an initial screening interview for "accentedness." Therefore, they may not have been representative of the population at large. Long argued that just because these cases were rare does not preclude them as potential test cases for the CPH, but severely limits any generalizations about typical adult second language abilities. It could be argued, however, that Neufeld (1979) was not trying to discount the hypothesis with one study. Rather, his position was that 'nonuniversality' constitutes the principle flaw in the maturational constraints argument" (p. 236). Although Long concluded that Neufeld's findings do not constitute counterevidence to the idea that there is a sensitive period for second language acquisition, one must recall that the CPH states that no one who commences learning a second language as an adult will be able to speak the L2 without a foreign accent, whereas Neufeld has shown that it may be rare but achievable.

Amount of native language use. A variable related to one's ability to produce target language sounds is the amount of native-language use. Flege et al. (1997) examined the effect of L1 use on production of an L2. Two groups of native Italian subjects who immigrated to Canada between the ages of 2.6 and 9.6 years and whose average residence in Canada was 18 years were used. The groups varied according to their self-reported daily use of Italian. The participants read and recorded sentences in English that were later rated for degree of foreign accent by native English speakers. The

researchers found that, generally, the more individuals spoke their native language, the stronger their accent in their second language (Flege et al., 1997).

Guion et al. (2000) attempted to replicate the findings of Flege et al. (1997). This time, Spanish sentences recorded by bilingual Spanish-Quichua speakers and monolingual Spanish speakers were presented to native Spanish listeners. Likewise, Quichua sentences produced by Quichua-Spanish bilinguals and near-monolingual Quichua speakers were presented to near-monolingual Quichua listeners. In both cases, the listeners were instructed to rate degree of foreign accent; in the case of the Spanish sentences, they were asked to rate Quichua accent, and in the case of the Quichua sentences, they were asked to rate Spanish accent. As in the Flege et al. (1997) study, the more the L1 was reportedly used, the greater the perceived foreign accent in the L2. Moreover, although L1 use was related to degree of accent in the L2, it was not related to L1 production. Therefore, an asymmetrical relationship exists between the L1 and the L2 sound systems that may have some bearing on the pronunciation of an L2, but the L2 sound system does not seem to be related to the production of the L1 (Guion et al., 2000). This evidence is counter to Grosjean's argument which claims that the influence of the L1 and the L2 is bi-directional.

Flege et al. (1997) argued, as did Grosjean (1992), that bilinguals' degree of L1 activation, or how much the L1 is used on a daily basis, influences their L2 production accuracy, rather than only neurological maturation at the time L2 learning commences. This single system hypothesis, as discussed earlier, contends that bilinguals have a single phonological system in which sounds from their L1 and L2 reside, making it difficult for them to isolate either system fully.

Experience. The amount of experience an individual has with the second language has also been found to relate to the speaker's production and perception of the L2. Bohn and Flege (1990) found that the perception of English vowels by adult native speakers of German improved somewhat linearly with the length of time spent in an English-speaking environment. For example, L2 experience did not affect perception of vowels that were similar in German and English. However, for the English vowels that were new to native German speakers, the Germans more experienced with English more closely resembled the native English speakers than did the inexperienced Germans. Likewise, Flege, Bohn, and Jang (1997) found that, as the amount of experience with the L2 increased, participants produced and perceived English vowels more like natives. In this case, "experience" was defined as the length of residence in the U.S. In both of the above studies, production and perception varied as a function of the relationship between the native language and English phonology. Although Flege (1988) found that the amount of second-language experience is related to participants' ability to detect a foreign accent, it was not related to adults' L2 production accuracy. Rather, it was the age of learning that was found to be more strongly related to the degree of perceived foreign accent.

Speaker and Listener Variables

So far, only speaker characteristics have been considered as they relate to the degree of perceived foreign accent. However, some factors believed to contribute to the amount of perceived foreign accent are present in both the speaker and the listener and seem to interact. These factors are the speaker and listener relationship, and familiarity with various aspects of the language and message.

Wijngaarden, Steeneken, and Houtgast (2002) also found that the degree of perceived foreign accent decreased with experience with the second language, and determined that a shared L1 facilitated intelligibility for nonnative speakers. They examined the intelligibility of Dutch, English, and German sentences produced by native and non-native speakers for trilingual listeners. Listeners in this study differed in their amount of experience with and proficiency of German and English. Specifically, the listener groups differed in their experience with German, with one group reporting weekly use and the other group reporting only yearly use. German proficiency differed between the listener groups as well; although both groups were fairly equal in English proficiency, the group who spoke German more often was more proficient than the group who rarely spoke German. To measure sentence intelligibility, the SRT method was used which is an adaptive method that measures the speech-to-noise ratio at which 50% of the tested sentences are perceived correctly. For this study, after the presentation of each sentence, listeners orally responds by repeating the sentence to the experimenter. The listeners who were highly proficient in English found English sentences spoken by native German speakers less intelligible than those spoken by native English speakers. For the same listeners, who were less proficient in German, the German sentences were found to be more intelligible when produced by non-native German speakers versus native talkers. The authors concluded that highly proficient listeners were able to use subtle phonetic cues present in native speech, for the less proficient listeners, these cues were not as helpful. This was because the less proficient L2 listeners were not able to categorize the L2 phonemes as natives. Instead, they perceived L2 speech as more intelligible if the sounds were matched to their L1 phonemes, as would be done by non-native speakers of

their same L1. This ability to categorize the L2 phonemes like native speakers increases with the amount of experience one has with the target language (Bohn & Flege, 1990; Flege, Bohn, & Jang, 1997).

Bent and Bradlow (2003) also investigated the relationship between speakers' and listeners' native language background and speech intelligibility. They found that a shared native language between a speaker and a listener facilitated speech intelligibility. In their study, native Chinese, Korean, and English speakers recorded simple English sentences for presentation to listeners from those same L1s and other native language backgrounds. The listeners' task was to listen to the sentence and write down whatever she or he heard. In general, non-native listeners found native speakers to be more intelligible than speakers from other first language backgrounds. Interestingly, for non-native listeners, the speech of non-native speakers from the same L1 was found to be as intelligible as that of native speakers. The authors' conclusion was consistent with that of Wijngaarden et al. (2002): non-native speech perception is associated with the relationship of shared speaker and listener L1. L2 speakers with the same native language share linguistic knowledge of both the L1 and the L2. In contrast, L2 speakers who differ in native language background share only linguistic knowledge of their target language. Therefore, a non-native listener is better able to interpret the L2 speech of a speaker with the same L1 compared to a speaker with a different L1, even if the speech differs greatly from the target language norm.

Characteristics of the stimuli and scaling methods

Thus far, speaker and listener variables have been considered as they influence the degree of perceived foreign accent and production in an L2. However, characteristics

beyond those of the speaker and listener influence the ease with which one is able to produce a second language and the degree to which one is judged to have a foreign accent. Specifically, there exist certain characteristics of the speech signal and differences in the types of scales used that influence the degree of perceived foreign accent.

Temporal and Acoustic Properties of Stimuli. The kind of stimuli used no doubt affects the degree to which accent is perceived and the way in which a second language is produced. Gottfried and Beddor (1988) presented spectral and temporal manipulations of French vowels to French and English listeners who were asked to listen to and identify the vowel in each syllable. The vowels were identified differently by the two language groups. In contrast to the French listeners, native English listeners were influenced by vowel duration in their categorization of the vowels, rather than only spectral cues, which is consistent with the prominent role of duration in the English vowel system. The researchers concluded that how one perceives a given vowel contrast in terms of spectral or temporal cues does not simply follow from experience with those vowels through their use. Rather, how one perceives a foreign language sound depends on the extent to which its acoustic properties correlate within one's L1 phonetic system. This argument is consistent with the single system hypothesis proposed by Flege and Eefting (1987).

Speaking Rate. Manipulations of various aspects of the speech signal have been shown to affect listeners' evaluations of native and foreign accented speech in somewhat counterintuitive ways. Increasing the speaking rate, for example, resulted in more native-like accent ratings of Mandarin speakers (Munro & Derwing, 1998). The ideal rate of nonnative speech was found to be somewhat slower than that of native speech, but faster

than what nonnative speakers typically produce (Anderson & Kohler, 1988). Munro and Derwing (2001) found that there is a point beyond which an increase in rate decreases accent ratings.

Length of Stimuli. Although speaking rate certainly affects the accent ratings of foreign speech, other manipulations, such as duration of the stimuli, have yielded similarly interesting results. Flege (1984) isolated progressively shorter units of English speech to determine whether a French foreign accent was detectable in English. In these experiments, listeners simply had to identify each item as having been spoken by a native or non-native speaker. First, he used sentence-length stimuli and found that listeners were able to detect a French accent. Then, he digitally manipulated the stimuli, isolating only the first word from the sentence and then only the first syllable from an utterance and, again, both trained and untrained listeners were able to detect a French accent. Finally, Flege found that even in short bursts of 30 ms, obtained from the digitally edited /tu/ syllable, native English listeners were able to detect the presence of a French accent.

Rating Scales. Different types of scales have been used in experiments involving the rating of foreign accent (Flege, 1988; Flege, 1995, Flege, Yeni-Komshian & Liu, 1999; Guion & Flege, 2000; Magen, 1998; Meador et al., 2000; Munro & Derwing, 1998, 2001; Piske, MacKay, & Flege, 2001; Southwood & Flege, 1999). In fact, Flege has spoken of the appropriateness and benefits of using an equal appearing interval scale (EAI) for the rating of foreign accents. He compared it to direct magnitude estimation and concluded that the EAI scale was preferred. Further, he has argued that the range should be nine points. However, he never explicitly explained why it is a better choice than the direct magnitude. A personal communication with James Flege (May 27, 2004)

revealed his reasons for using this type of scale rather than the direct magnitude estimation:

"Parsimony. How many different degrees of accentedness can listeners reliably discern? In the Southwood/Flege paper, we provide preliminary evidence that a seven-point scale under utilizes listeners' ability, whereas a nine- point scale does not. Of course, as you will have noted in the Piske et al. paper that a nine point scale and a continuous scale yield much the same results."

The use of the EAI scale is also consistent with the rating scales used in much of the foreign accent literature. A review of the methodologies of studies that had listeners rate the degree of foreign accent they perceived in speakers revealed that, with the exception of one (Flege, 1988), all used equal appearing interval scales (Flege, Yeni-Komshian & Liu, 1999; Guionet al., 2000; Magan, 1998; Munro & Derwing, 1998, 2001; Piske, MacKay, & Flege, 2001; Southwood & Flege, 1999). In fact, two of them specifically discussed the methodologies in terms of type of rating scale and concluded that listeners were able to partition L2 foreign accents into equal intervals, so it is appropriate to use an EAI scale in foreign accent studies. Other researchers have examined various types of scales and the results support the use of the EAI scale.

With regard to the range of the scale vales that are needed to exploit listeners' full range of sensitivity, Southwood and Flege (1999) found that "a nine-point rating scale should be used to rate L2 speech samples for degree of foreign accent" (Piske, MacKay, & Flege, 2001, p. 195). This happens because, when rating foreign accents, a potential ceiling effect could occur due to the number of scale intervals used. Southwood and Flege noted that, although seven-point scales are frequently used, they may not be

sensitive enough for all listeners to discriminate among the stimuli, in this case, the speakers' sentences. Additionally, these authors examined the use of the five- and seven- and nine-point scales for native and nonnative sentences. The five-point scale failed to yield a significant between-group difference, whereas the nine-point scale did.

Linguistic variables

Additional factors specific to the languages under investigation also contribute to a speaker's foreign language production and degree of foreign accent. As Bent and Bradlow (2003) noted, there are specific linguistic contributions of the L1 during the production of the L2 which serve to mediate intelligibility between L2 speakers and listeners. Some of these specific interactions are discussed next.

Phonological similarities and differences of the L1 and the L2. Specific and unique interactions exist between the L1 and L2 phonological systems that may predict areas of pronunciation difficulty for the second language learner (Bohn & Flege, 1990; Flege & Eefting, 1987; Guion et al., 2000; Munro & Derwing, 1998). To better elucidate the difficulties encountered by native Spanish speakers learning English, a discussion follows that explores the unique contributions of the Spanish and English phonological systems to pronunciation. This discussion is relevant to the present study.

The Spanish and English phonological systems have specific differences that may create difficulty for a native Spanish speaker learning English. Only the more salient and those most relevant to the present study are discussed. It should be noted that the following observations are generalizations and that not all Spanish and English dialects exhibit these characteristics.

The differences between the English and Spanish phonological systems include syllable structure, vowel quality, consonants, and stress. One of the factors affecting syllable structure is the insertion of the initial and non-initial epenthetic /ə/ by Spanish speakers before an English cluster (Magan, 1998). For example, Spanish speakers may say /əstaemp/ instead of /stamp/ because the Spanish language has different clusters than English. The insertion of the epenthetic schwa facilitates production of the English clusters by breaking the cluster into two syllables: /əs/ and /taemp/.

Vowels are another area in which Spanish and English differ. Whereas English has 14 vowels, Spanish has only five. English speakers tend to reduce the vowels in unstressed syllables, whereas Spanish speakers are more likely to produce them fully, e.g., *seas*[o]*ns* for *seas*[ə]ns. Also, in Spanish accented speech, English lax vowels tend to be produced as tense vowels, e.g., *ch*[i]*p* for *ch*[I]*p* (Magen, 1998). Even during rapid speech, vowel length is maintained, and Spanish speakers tend to delete syllables or consonants rather than shorten the vowel (Iglesias & Anderson, 1993).

Spanish and English also have different consonants and phonological rules. Spanish speakers tend to drop word final /s/ or /z/ to simplify final clusters not allowed in Spanish. In Spanish, plurality is marked by redundancy across the verb phrase, so one can delete the final /s/ and still convey plurality by marking it in the verb (Iglesias & Anderson, 1993). Spanish speakers learning English may therefore say *I saw three cat* instead of *I saw three cats.* Also, Spanish does not indicate possession by using an apostrophe /s/, /z/ or / z/, so, instead of saying *the girl's dog* Spanish speakers would say *the dog of the girl*. So an English sentence such as *I saw the girl's dog* may be produced as *I saw the girl dog*. Spanish has only one affricate, /tʃ/, which occurs in word-initial

and intervocalic positions and is commonly substituted for the English fricative /ʃ/. There are voicing distinctions between intervocalic /s/ and /z/ in English that are typically not produced by Spanish speakers, leading to substitutions such as *free*[s]*er* for *free*[z]*er*.

Finally, there are subtle stress pattern differences between English and Spanish. In terms of lexical stress, Spanish typically places the stress on the final syllable in multisyllabic words, whereas English stress patterns call for the stress on the penultimate or antepenultimate syllable. Therefore, Spanish speakers tend to stress the final syllable of English multisyllabic words, e.g. *combinatíon* for *combinátion*. Also, Spanish has two degrees of stress (weak and strong), whereas English has three (strong, medial, and weak) (Stockwell & Bowen, 1965). Further, the vowel qualities differ between those found in weak-stress syllables versus those found in syllables with greater stress. In English, word-level stress patterns can be used to differentiate nouns from verbs, such as in *ínsult* and *insúlt* (Stockwell & Bowen, 1965). In Spanish however, word stress is less varied and of different kinds. Verb endings such as *o*, first person singular, and *ó*, third singular past, exemplify stress contrasts that change meaning in Spanish. For example, *háblo* means *I am speaking*, while *habló* means *he spoke* (Stockwell & Bowen, 1965). Finally, English has a relatively regular alteration between syllables of weak stress and those of strong stress, whereas Spanish has fairly long sequences of weak-stress syllables. This can be seen in the differences in the stress patterns between English *òperátion* and Spanish *operación* (Stockwell & Bowen, 1965).

The present study attempted to determine whether increased phonological difficulty, defined by the number of target phonemes deemed most difficult for Spanish

speakers learning English, would be related to an increase in perceived foreign accent. Manipulation of this variable is unexplored in the literature thus far.

Purpose of the Present Study

Despite the differential outcomes of various studies on age-related language learning constraints, several generalizations can be made (Long, 1990). Adults proceed through *early* stages of morphological and syntactic development more quickly than children. This differential rate of acquisition may be the result of older learners' more advanced metalinguistic skills. In other words, adults may have a better ability to think about language and use for their learning of an L2 some of the same skills they acquired in learning and mastering their L1.

So far, studies have been conducted to determine whether the amount of language use affects the degree of perceived foreign accent by Italian and Canadian-English speakers (Flege et al., 1997), Chinese and English speakers (Flege, 1988) and Spanish and Quichua speakers (Guion et al., 2000). In support of the ECH, it has been shown that the more individuals use their L2 the better able they are at gauging the degree of foreign accent in their target language. Perhaps, as asserted by the ECH, when compared to adults who use their second language rarely, adults who often use their second language establish phonetic categories in their target language that more closely approximate those of a native speaker, making them better able to detect foreign accents in their L2 than individuals who use their L2 less often. The first aim of the present study was to extend these findings to Spanish-English bilingual speakers.

In the present study, the question of whether the degree of perceived foreign accent in an L2 is related to the listener's amount of L1 use was also investigated. Bohn

and Flege (1990) found that the perception of English vowels by adult native speakers of German improves with their amount of experience with English. Likewise, Flege, Bohn, and Jang (1997) found that as the amount of experience with the L2 increased, subjects perceived English vowels more like natives. In these studies, experience was defined as the length of residence in the U.S. Additionally, Best and Bradlow (2003) and Wijngaarden et al. (2002) found that intelligibility of L2 speech improves if the listeners had shared L1s. Although those studies focused on the amount of L2 exposure, the present study investigated the amount of L1 use. In the present study, the relationship between amount of daily L1 use and perceived degree of foreign accent were explored. It was hypothesized that individuals who used their L1 less would be better able to detect an accent in nonnative speakers of their L2 than would individuals who use their L1 more often.

The second purpose of the present study was to determine whether the phonological difficulty of the stimuli would affect how well the speaker was able to produce them and whether this increased phonological difficulty would result in a more detectable accent. Difficulty was determined by how many potentially challenging targets were included in the stimuli. It was predicted that the level of phonological difficulty of the stimuli would affect the degree of perceived foreign accent, with more difficult sentences eliciting higher accent ratings. The less difficult sentences were predicted to be produced with less of an accent and, therefore, rated as less accented by all listener groups. The more difficult sentences were predicted to be produced with more of an accent and be rated as more accented by all listener groups. The low L1 use group was predicted to rate sentences as more accented than would the high L1 use

group. This was predicted because the low L1 use group would have more experience with English and would therefore, more readily notice deviations from native-language norms in the productions of the speakers.

Chapter 2

Method

Design

The experimental design was a 3 x 3 x 2 mixed-model factorial. Listener group was varied between subjects (native English, high-L1 use bilingual, low L1-use bilingual), and both speaker category (heavily accented bilingual, moderately accented bilingual, and slightly accented bilingual) and level of phonological difficulty (less difficult, more difficult) were manipulated within subjects. The dependent variable was ratings of accentedness.

Participants

Speakers. Three groups were used as speakers: one heavily accented bilingual Spanish-English speaking group, one moderately accented bilingual Spanish-English speaking group, and one slightly accented bilingual Spanish-English speaking group, with two speakers per group. These participants were recruited from the University of South Florida English Language Institute (ELI) and through advertisements within the Departments of Psychology and Communication Sciences and Disorders (CSD) (see Appendix A). Extra credit points were offered to those subjects enrolled in courses within the Psychology and CSD Departments.

The inclusion criteria were that speakers were Spanish/English bilingual women between the ages of 18 and 35 years who did not report a history of hearing, speech, or language disorders. The investigator used her clinical judgement to discern whether any speakers had any speech defects that would preclude them from being participants. Race was not considered as either an inclusion or exclusion criterion. According to self-report, all participants had Spanish as their L1 and English as their L2. Additionally, these participants had not learned an L3 because, according to the equivalence classification hypothesis, they may have experienced interference from their third language that would have affected their first and second languages in ways that were dissimilar from bilingual speakers. As shown in Table 1, the speakers in each group were of roughly the same age and came from five different Spanish-speaking countries.

Table 1. Speaker Characteristics

Speaker Group	Mean Age in Years	Country of Origin
Slightly Accented	26.5	U.S., Dominican Republic
Moderately Accented	25.5	Dominican Republic, Colombia
Heavily Accented	27.0	Venezuela, Mexico

Listeners. A power analysis was performed to estimate a sample size that would ensure a power of .82. Using the Pearson-Hartley charts (Meyers & Well, 1995), it was determined that 12 subjects per group would be needed to obtain the desired significance level of .05. Twelve native monolingual English speakers from the United States comprised the first group. The listeners in the other groups were 24 native-Spanish speakers, 12 higher use and 12 lower-use, who learned English as a second language

some time after adolescence and spoke no other languages. All listeners were between the ages of 18 and 45, and the groups included both males and females. Listeners were recruited from the English Language Institute and the Departments of Psychology and Communication Disorders at the University of South Florida and were offered extra credit points for their participation (see Appendix A), paid $5.00 cash, or provided the equivalent in gift certificates to on-campus restaurants. They all passed the Speech Listening Test (Griffiths, 1967) before being accepted into the study, indicating that they had normal hearing and speech discrimination abilities. This test demonstrated the listeners' ability to discriminate speech sounds presented in groups of words that differed in initial or final consonant, such as *lake, rake take, bake*.

These individuals also completed a language background questionnaire (see Materials) and, based on their responses, were designated as either higher-use Spanish L1 listeners or lower-use Spanish L1 listeners. Participants who indicated that they spoke Spanish 50% of the time or more were placed in the high-L1 use group, and those who indicated that they spoke Spanish 49% of the time or less were place in the low-L1 use group. As shown in Table 2, the low-use listeners' mean amount of daily L1 use was 23% and the high-use listeners' daily amount of L1 use was 72.1%. Although all three groups of listeners were fairly similar in terms of gender, there was significant variability in terms of their country of origin, with 12 countries being represented among the 24 bilingual listeners.

Table 2. Listener Characteristics

Listener Group	Gender	Mean Age In Years	County of Origin	Mean Percent Self-reported Daily L1 Use
Native English	M 5 F 7	29.4 (8.9)	U.S.	
Low-use Bilingual	M 5 F 7	22.2 (3.3)	Colombia, Cuba, Nicaragua, Dominican Republic Guatemala, Spain, Venezuela, El Salvador, Puerto Rico	23% (14.86)
High-use Bilingual	M 4 F 8	27.9 (8.1)	Colombia, Mexico, Peru, Cuba, Venezuela, Puerto Rico, Panama	72.1% (8.38)

M=Male; F=Female Standard deviations are presented in parentheses.

Materials

Demographic questionnaire for the speakers. Potential speakers completed a questionnaire prior to being accepted in the study. The questionnaire was reviewed before the speaking task began to ascertain whether the individual met the inclusion criteria. It consisted of the following:

1. What is your country of origin?
2. At what age did you begin to learn English?
3. Do you speak any other language besides Spanish and English?
4. Have you ever been diagnosed with a speech, language, or hearing disorder?

In addition to the questions listed, participants also provided basic biographical information such as age, gender, and educational level.

Demographic questionnaire for the listeners. Responses to these questions indicated the listener's language preference and use in given social situations and, therefore, provided a measure of her overall daily L1 use. The questionnaire was reviewed during

the listening task, since it was not necessary to ascertain in which group the listener would be included prior to this.

1. At what age did you begin to learn English?
2. Which language do you use most at home?
3. Which language do you use most at work?
4. Which language do you use most at parties?
5. Which language do you use most with friends?
6. On average, how much do you use Spanish daily?
7. If you learned in school, how old were you when you first started English classes?
8. How many times/hours per week did you have English class at that time?
9. Do you or have you ever had a diagnosed speech and/or language disorder?
10. Do you speak any other language besides English or Spanish?

In addition, participants provided basic biographical information such as age, gender, birthplace, and educational level.

Questions two, three, and five were taken from a previous study, which also looked at the effects of L1 use on the degree of foreign accent (Guion et al., 2000). Questions four and five are a modification and combination of several of the questions used in that same study (Guion et al., 2000). Those authors had asked their participants about language usage with siblings, with friends at parties, and when meeting friends on the street. In order to participate in the study, individuals indicated in question 1 that they learned English after adolescence (operationalized as 12 years of age), and in question 10 they indicated that they spoke no other languages besides English and Spanish.

Questions two and three were used to determine the participants' preferred language for family and work situations because these areas constitute the bulk of their spoken interactions. Questions four and five provided an indication of the participants' social language use. Question six was included to obtain the participants' overall estimation of their daily L1 use. Individuals were assigned to higher-use group if they indicated that Spanish was the main language used at work or home (questions 2 and 3) and indicated that Spanish was the main language used in social settings (questions 4 and 5). Additionally, to be included in the higher-use L1 group, individuals noted that they used Spanish more than half the time on a daily basis. Questions 7-9 were included for descriptive purposes only.

Practice sentences. Prior to the speaking and listening tasks, participants had the opportunity to practice the task using ten sentences. These sentences are listed in Appendix B and were constructed to be simple enough for speakers to produce and listeners to comprehend with little difficulty.

The speakers for the listening practice trials were different from those in the actual listening task and included bilingual speakers judged to have a both strong and slight Spanish accents by doctoral students and faculty in the Department of Communication Sciences and Disorders trained in phonology. Although the task was fairly simple and five practice sentences would probably have trained the listeners to the task, it was determined that using ten sentences would cause the listeners to focus on the accents of the speakers rather than on the meaning of the sentences. The use of anchors was therefore not necessary.

Stimulus recording. Because they were intended to elicit a foreign accent from the speakers, the stimuli for this study consisted of 20 English sentences, each containing several words with sounds presumed to be difficult for native Spanish speakers learning English. The stimuli were modified from those used in a previous study (Magen, 1998) and included phonemes and phrases in American English that were expected to result in the production of a foreign accent when spoken by Spanish speakers. Specifically, they were constructed to include phonemes and phoneme sequences that are present in English but not in Spanish or that differ between English and Spanish (Magen, 1998; see Appendix C). Sentences, rather than words, were used because they more nearly approximate the short utterances typical of conversational speech.

The sentences were grouped as either phonologically more difficult or phonologically less difficult based on the number of "targets" (i.e., challenging phonemes or phonemic sequences) present in each sentence. The more phonologically difficult sentences contained between four and eight speech targets. The less phonologically difficult sentences contained at most three speech targets. The number of targets per sentence was found to be significantly different between the levels, ($t(18)= 2.75$, $p< .05$). It was predicted that the more phonologically difficult sentences would result in a greater number of errors from L2 native-speaker norms by the heavily accented group relative to the slightly accented speaking group. It was expected that the differences from L2 native-speaker norms that the speakers made would vary according to their level of accent, with heavily accented individuals producing more errors on the targets than the slightly and moderately accented individuals. It should be noted that although the term "error" is used throughout this paper, the authors recognize that non-native speakers are

really producing differences compared to native speaker norms and not error per se. It was expected that the speakers would insert an epenthetic /ə/ before fricative plus stop clusters in English syllables that are not allowable in Spanish. It was also expected that the speakers would produce the vowels in the unstressed syllables fully, as tense vowels. The Spanish speakers were also expected to drop the word final /s/ or /z/ to simplify final clusters not allowed in Spanish, and because plurality is marked by redundancy across the verb phrase, the plural /s/ was expected to be dropped. The speakers were also expected to substitute their only affricate, /tʃ/, for the English /ʃ/. The voicing distinctions between intervocalic /s/ and /z/ in English words were not expected to be produced by speakers. Finally, the speakers were expected to stress the final syllable of English multisyllabic words instead of the penultimate or antepenultimate as appropriate. Speakers were also expected to produce several weak stress syllables in a row, because unlike English, which has a relatively regular alteration between syllables of weak stress and those of strong stress, Spanish has fairly long sequences of relatively weak-stress syllables.

Recording of stimulus sentences. The speaking task was administered individually in the Acoustics Laboratory in the Department of Communication Sciences and Disorders at the University of South Florida. Upon arrival in the lab, speakers were greeted and told about the nature of the study. They were directed to have a seat and then were given the informed consent forms. Then, they were given the language background questionnaire and instructed to answer the questions to the best of their ability and recollection. After each participant completed the language background questionnaire, the investigator quickly determined whether she met the inclusion criteria before proceeding. All participants met the criteria and were provided the list of practice

sentences (see Appendix B) to rehearse the recording task. Approximately three minutes before the scheduled recording the speakers were given the stimulus sentences (see Appendix C) so as to allow for familiarization with the stimuli and ensure fluent speech during recording.

These sentences were provided on a sheet of paper with size 16 font so as to make them large enough to read easily. The speakers were instructed to read the sentences three times each at a normal conversational pace and to leave about three seconds between sentences. To avoid ambiguity, the experimenter demonstrated an acceptable pace by reading a practice, non-stimulus sentence to the speaker. For complete speaking task instructions, see Appendix D.

Once the speakers demonstrated understanding of the task and had familiarized themselves with the stimuli, the speaking task began. The experimenter exited the booth and returned to the recording equipment to monitor the experiment while the speaker read the practice sentences. After reading all 20 stimuli sentences three times each, the participant had completed the speaking task and was thanked for her time. The speaking task took approximately 20 minutes to complete.

A digital file was constructed by recording all three readings of the 20 sentences by each of the six speakers using a Roland VS-1824 24 bit digital studio workstation with a sample rate of 44.1kHz. After all the speakers recorded the sentences, speakers were assigned to the heavily accented group, moderately accented group, or the slightly accented group based on accent judgements made by doctoral students and faculty in the Department of Communication Sciences and Disorders trained in Spanish-English phonology. These individuals met and jointly listened to the speakers' recordings and

independently rated their accents using a nine-point rating scale. There was complete agreement among the listeners as to what group to assign the speakers. The two speakers who were categorized as slightly accented obtained a mean accent rating of 2.09. The two speakers placed in the moderately accented group received a mean accent rating of 5.08 and the two speakers in the heavily accented group received a mean accent rating of 8.76.

Stimulus preparation. The sentences produced by the speakers were saved in the workstation and to a disk for subsequent digital editing. They were edited using the digital waveform editing program Praat™ on a Dell computer so that only the second of the three sentences was used for presentation to the listeners. This method allowed the experimenter the option of using the first or third sentence in the series if the second was in some way inaccurately produced (e.g., hesitation or stuttering on a word or syllable). The interest was in the second sentence because it was thought that after having read the sentence once, the second reading would be produced more fluently. The selected sentence was saved in a separate file for each of the 20 sentences for all six speakers. Using the program Resample, the sentences were then resampled to a rate of 48,828 Hz to make it compatible for use with the other programs needed to create the experiment. The 120 files (six speakers x twenty sentences) containing the sentences were then arranged for order of presentation and the experiment was created and run using the ECoS Version 2™: Experiment Generator and Controller program. To avoid order effects, four different orders of the stimuli were created in which the sentences were randomized with the constraints that no more than two sentences per speaker would occur in succession. Sentences were presented to the listeners with approximately three

seconds between trials. To test for reliability, the first 20 sentences were presented again in the same order at the end of the experiment, but were not used for other statistical analysis.

Procedure

Participants in the listening task were tested individually or in groups of two. Upon arrival in the lab, they were greeted and told about the nature of the study. After giving consent to participate, they completed the language background questionnaire. The questionnaire was administered orally by the experimenter. This was done to better ensure that the participants understood the questions and that the experimenter understood with the listener's language background.

After it was determined that each individual met the criteria to participate, the listening task began. The listening task took place in the Acoustics Laboratory in the Department of Communication Sciences and Disorders. Participants were directed to have a seat in front of one of the computers. First, they completed the Speech Listening Test (Griffiths, 1967). This test consisted of 25 sets of four words that were minimal pairs. The minimal pairs selected did not include phonemes deemed difficult for Spanish speakers learning English in order to avoid interference from the listener's L1. Thus, they provided a clear picture of participants' speech discrimination abilities in the absence of L1 interference. In this test, the listener first had four practice items to familiarize him or her with the task and to allow for appropriate volume adjustment. The listener heard a word spoken via headphones and saw four words displayed on the computer screen (e.g. rake, take, make, bake). He or she was instructed to click with the mouse on the word that he or she thought was spoken. The test was scored after the

entire listening task had been completed. All participants scored 100% on this task, which indicated that they had speech discrimination abilities that were within normal limits and were included in the study. After the Speech Listening Test had been completed, the participants were invited to stand and take a short break before proceeding to the listening task. After the break, they were instructed to return to their seats in front of the computer. They then heard the experimenter read the listening task instructions. For complete listening task instructions, see Appendix E.

The listeners were told that they would use headphones to hear sentences spoken by non-native English speakers and that they should rate each sentence independently. Listeners were instructed to estimate the degree of Spanish accent in each sentence by using a 9-point rating scale that was displayed on the computer screen. The scale had the labels: "slight Spanish accent" at the left side of the scale and "strong Spanish accent" at the far right side of the scale, but did not display numbers. Listeners were instructed to drag the curser to any one of the nine points along the scale to indicate the degree of Spanish accent present in each of the sentences.

Each session began with ten practice trials consisting of non-stimulus sentences (see Appendix B) to familiarize the listeners with the use of the scale and to attempt to shift the listeners' focus from the meaning of the sentences to the speakers' accents. Following these practice trials, the listener was given the opportunity to ask questions regarding the task. Feedback during the practice trials was restricted to the use of the scale, not the rating of the speakers' accents. For both the practice sentences and the stimulus sentences, the listener dragged the cursor to a point along the scale that corresponded to a number representing the degree of foreign accent that he or she judged

the speaker of the sentence to have. After each trial the participant clicked "Accept" at the top of the screen to advance to the next trial. After the listener had rated all sentences, he or she was asked whether any voices sounded more distinctive than others and whether this recognition influenced their ratings. At this point, the listener had completed the listening task and was free to leave.

Data Reduction

After the experiment was finished, the data were saved in an Excel file that contained the responses sorted by listener type, speaker type, and stimulus type. A mean rating was calculated for each speaker group and stimulus type, which yielded six values per listener. These data were put into SAS for statistical analysis.

Chapter 3

Results

This study investigated the relationship between amount of daily first language use, the phonological difficulty of the stimuli, and degree of perceived foreign accent. It was hypothesized that individuals who use their L1 less often, in this case the low-use bilinguals, would be more able to gauge the degree of foreign accent like native speakers compared to those who use their L1 more often when rating the degree of accent in nonnative speakers of their L2. The second purpose of the study was to determine whether the phonological difficulty of the stimuli would affect how well the speakers were able to produce sentences and whether this increased phonological difficulty would result in a more detectable accent. Difficulty was determined by how many potentially target phonemes were included in the stimuli. It was predicted that the level of phonological difficulty of the stimuli would affect the degree of perceived foreign accent, with the more difficult sentences eliciting higher accent ratings.

The study required that speakers who differed according to their degree of foreign accent (slight, moderate, and heavy) read English sentences that varied by level of phonological difficulty (more or less difficult). The sentences were then rated for degree of foreign accent by listeners who differed according to amount of daily L1 use but not necessarily in proficiency as this variable was not controlled.

Intra-rater Reliability

To determine intra-rater reliability, the first 20 sentences were presented again in the same order at the end of the experimental procedure. A Pearson Product-Moment Correlation was run to investigate intra-rater reliability for ratings of the repeated sentences. The correlation coefficients for the native English listener group ranged from .63 to .96 (M = .80, SD = .098). The correlation coefficients for the low-L1 use bilingual group ranged from .29 to .93 (M = .80, SD = .18). When the score of the outlier in the low-use bilingual group was removed, the mean was .86. A correlation of .30 is considered to be moderate, therefore, the decision was made to include the data of the outlier. Also, the mean correlation for the low-L1 use group was still higher than that of the high-L1 use group with the outlier included. The correlation coefficients for the high-L1 use bilingual group ranged from .48 to .93 (M = .75, SD = .15). With the exception of the one outlier, all of the correlation coefficients for all listening groups were considered moderate to high. However, there was greater variability among the listeners in the bilingual groups. Generally, it was determined that participants were able to do the task with an acceptable level of reliability.

Degree of Perceived Accent

A three-way mixed-model analysis of variance (ANOVA) was run analyzing speaker, listener, and sentence difficulty factors. Speaker group (strongly accented, moderately accented, and slightly accented native Spanish) was a within subjects factor, listener group (native English, high-L1 use bilingual, and low-L1 use bilingual) was a between subjects factor, and difficulty of sentences (more difficult and less difficult) was a within subjects factor. Comparisons were conducted for all significant effects related to

the hypotheses. All effect sizes were calculated using the formulas provided by Rosenthal and Rosnow (Levine & Hullett, 2002).

As predicted, there was a main effect of degree of speaker accent on perceptions of accentedness, $F(2, 66) = 317.25$, $p < 0.0001$, $\eta^2 = .91$. As shown in Figure 1, the slightly accented speaker group was rated as least accented ($M = 2.269$, $SD = 1.06$). The moderately accented speaker group received ratings intermediate to the slightly accented and heavily accented speaker groups ($M = 3.54$, $SD = 1.09$). The heavily accented speaker group was rated as most accented by all listener groups ($M = 6.82$, $SD = .92$).

Although there were no other main effects, there were two interactions. A significant two-way interaction was found between speaker and listener group, $F(4,66) = 3.61$, $p < .0159$, $\eta^2 = .22$. Tests of Least Significant Differences (LSD) revealed that the high L1-use bilingual listener group rated the slightly accented speaking group as significantly more accented than did the native English listener group (see Figure 1). There were no differences among listener groups in how they rated the moderately and heavily accented speaker groups. Thus, in gauging the degree of foreign accent of the slightly accented speaker group, it could be that amount of L1 use is inversely related to the ability to detect subtle accent or that the high L1 use group rated the slightly accented speaker group more harshly.

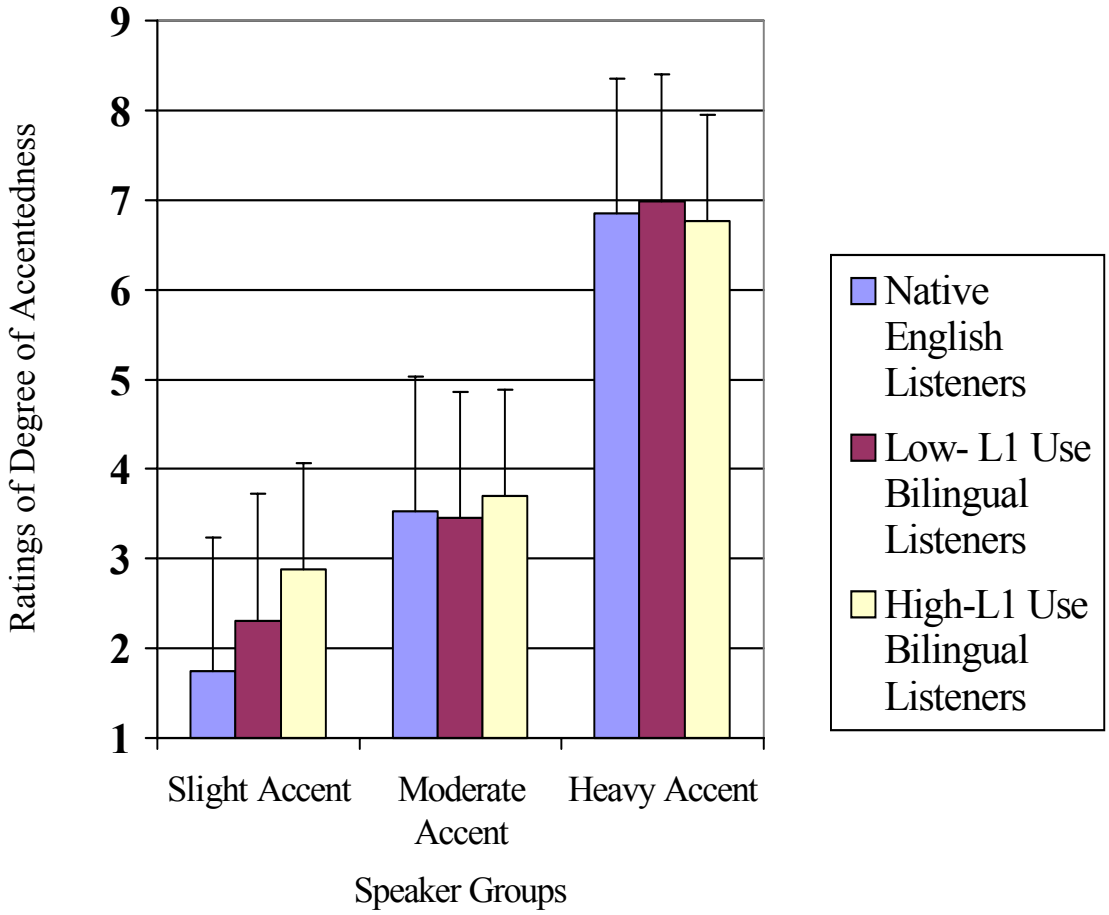
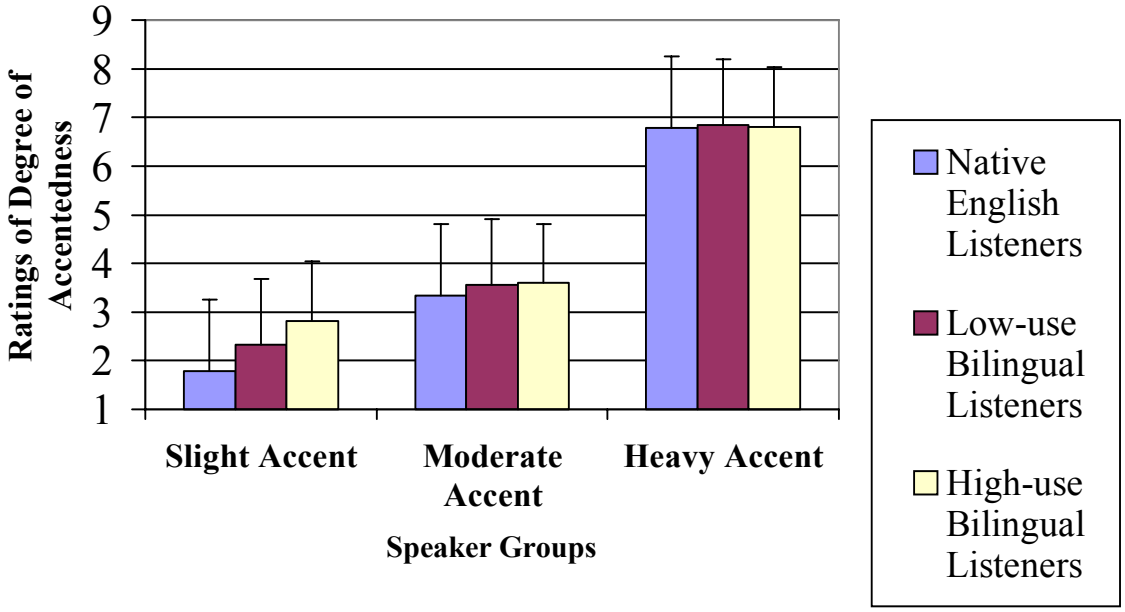


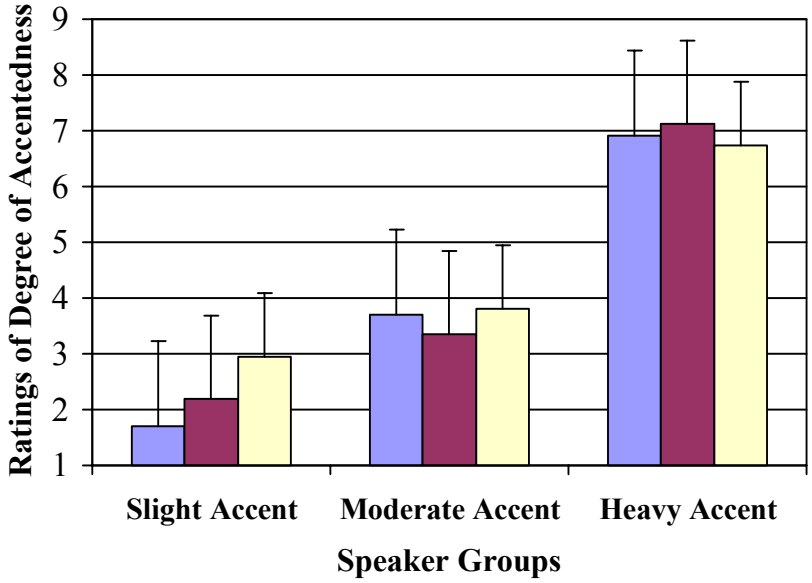
Figure 1. Mean Accentedness Ratings by Listener and Speaker Groups

A significant three-way interaction was found for speaker group, listener group, and phonological difficulty, $F(4, 66) = 2.47$, $p < .0530$, $\eta^2 = 0.15$. This interaction was not significant after the Greenhouse-Geisser Epsilon correction was made, and was significant only at p values of .0721. However, after checking the data for additivity and sphericity, it was concluded that the results were acceptable before the correction, and were therefore significant. As Figure 2 suggests, LSD tests revealed that the high L1-use

listener group rated the slightly accented speaking group as significantly more accented than did the native English listener group for both the more and less phonologically difficult sentences. The high L1-use listener group also rated the slightly accented speaker group as significantly more accented than did the low L1-use listener group for the more phonologically difficult sentences. There were no differences found between listener groups in how they rated the moderately and heavily accented speakers in terms of the phonological difficulty of the stimuli. Generally, the listening groups performed similarly, with the exception of their ratings of the slightly accented speaking group. Although these results were not expected, they are nonetheless compelling and require further attention and analysis.



Less Difficult Sentences



More Difficult Sentences

Figure 2. Mean Ratings of Accent by Both Speaker and Listener Groups for Both Less and More Difficult Sentences

Analysis of Individual Speakers' Ratings

In order to determine whether there were differences between the individuals within the speaker groups that may have attributed to the effects found among speaker groups, an analysis of variance (ANOVA) using a hierarchical design was run analyzing effects of speaker group, individual speaker nested in speaker group, listener group, and sentence difficulty factors on ratings of accentedness. There was a significant two-way interaction between individual speaker number and speaker group. As suggested by Figure 3, there were significant differences in perceived degree of accent between the two speakers in the heavily accented group, $F(2, 198) = 13.52$, $p < 0.0001$, $\eta^2 = .12$. All listeners rated these two speakers as the most heavily accented, with one speaker (Heavy 1) rated as more accented than the other (Heavy 2). Speaker Heavy 1 had a mean accent rating of 7.54 while speaker Heavy 2 had a mean accent rating of 6.2. Listeners tended to rate speakers in the other two groups similarly. In fact, both speakers in the moderately accented speaking group had a mean accent rating of 3.57. Generally, the results echoed those found for the analysis by speaker group.

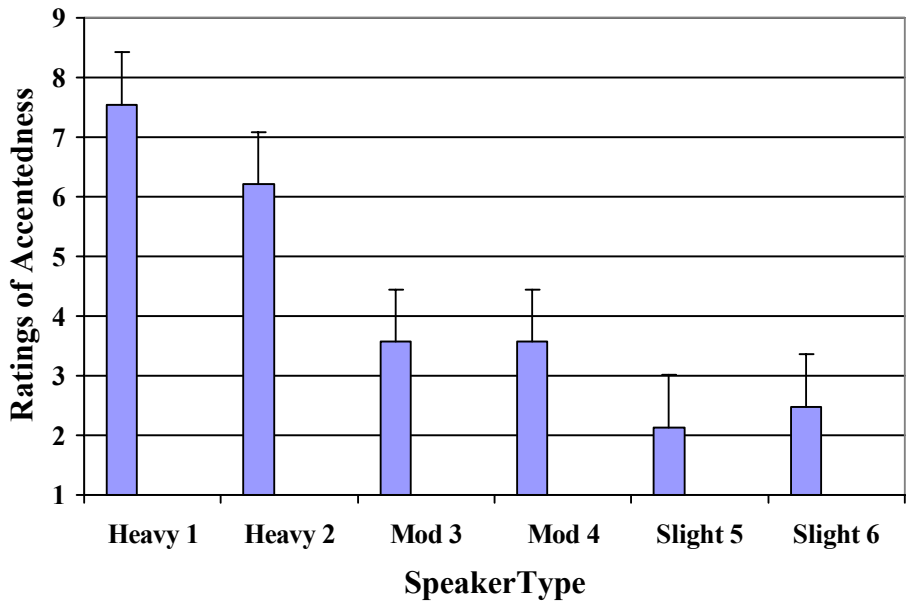


Figure 3. Mean Accentedness Ratings of Each Speaker

Another two-way interaction was found between speaker group and listener group, F (4, 198) = 2.636, $p < .035$, $\eta^2 = 05$. The high L1-use bilinguals rated the slightly accented group as significantly more accented than did the native English listener group. There was also a significant three-way interaction between listener group, speaker group, and difficulty level, $F(4, 198) = 3.376$, $p < .011$, $\eta^2 = .064$. Tests of Least Significant Differences revealed that the native English listener group rated the moderately accented speaker group as significantly more accented on the less difficult sentences compared to the more difficult sentences, see Figure 4. Also, the low L1-use bilingual listener group rated the heavily accented speaker group as significantly more accented on the more difficult sentences compared to the less difficult sentences. Finally, there was a three-way interaction found between speaker group, individual speaker, and difficulty level, F (2, 198) = 4.293, p < .0001, η_2=.173. As Figure 5 suggests, LSD tests revealed that

speaker 2 (heavily accented speaker 2) was rated as significantly more accented on the more difficult sentences, speaker 3 (moderately accented speaker 1) was rated as significantly more accented on the more difficult sentences, and speaker 4 (moderately accented speaker 2) was rated significantly more accented on the less difficult sentences.

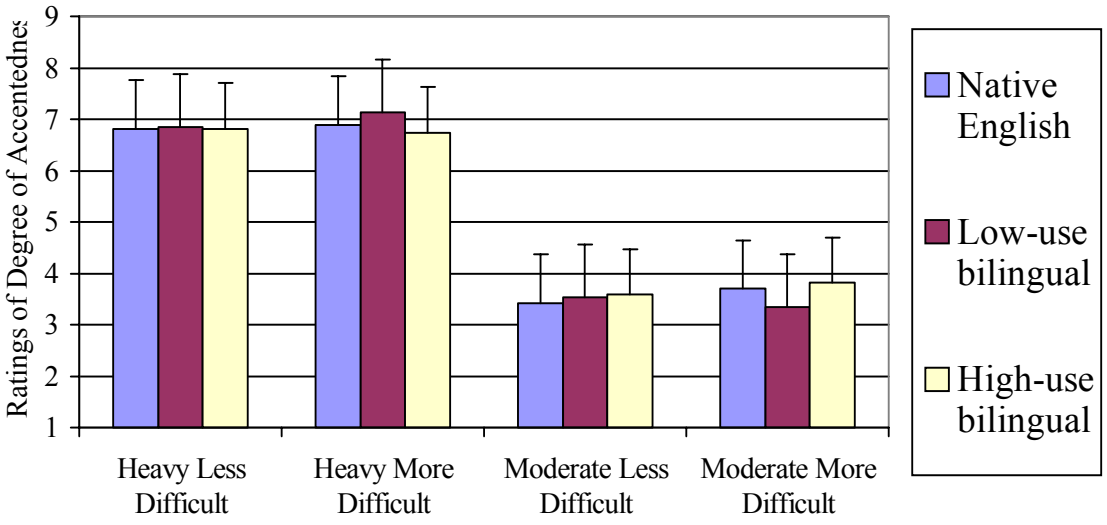


Figure 4. Mean Ratings by Listener and Heavily and Moderately Accented Speaker Group for Less and More Difficult Sentences

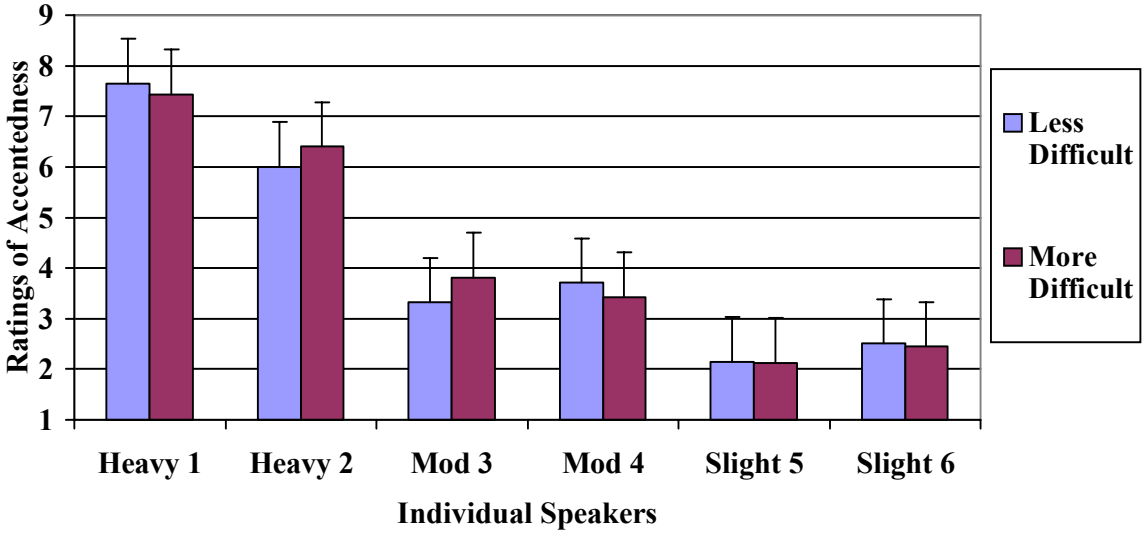


Figure 5. Mean Accent Ratings for Individual Speakers for Less and More Difficult Sentences

Analysis of Speakers' Production of Target Stimuli

The analysis of production errors was begun by calculating the percentage of targets correct in each sentence produced by each speaker. In order to determine whether the speaker had produced the target differently from native speaker norms, the investigator and an independent rater listened to the recordings over headphones using the digital waveform editing program Praat™, which also provides a visual display of the speech samples showing both waveform and spectrogram. The targets were then marked as correct or incorrect productions and tallied. An independent rater listened to the entire sample of the more heavily accented speaker (Heavy 1) and 10% of the sentences of the remaining speakers. Agreement between the two raters was 79.76% for the heavily accented speaker's sentences and 88.10% for the remaining speakers' samples. In cases of disagreement, the raters conferred until agreement was reached. Total differences in production compared to native norms were obtained by sentence (see Figure 11). The

slightly accented speakers produced the fewest errors. The moderately accented group produced more errors, although they differed only somewhat from the slightly accented speaker. The heavily accented speaker group produced the most target errors for both the less and more difficult sentences. The analysis by individual speaker yielded similar results as the analysis by speaker group.

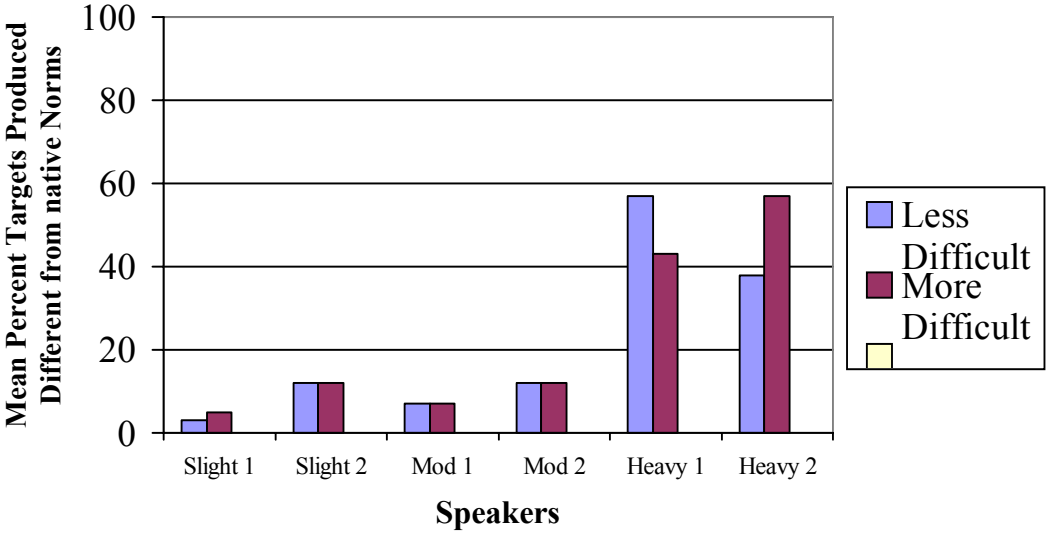


Figure 6. Mean Percent Target Errors by Individual Speakers and Sentence Type

To further investigate the number and types of differences from native norms produced by each speaker, the percentage of correct productions for each target were calculated for each target for the less and more difficult sentences.

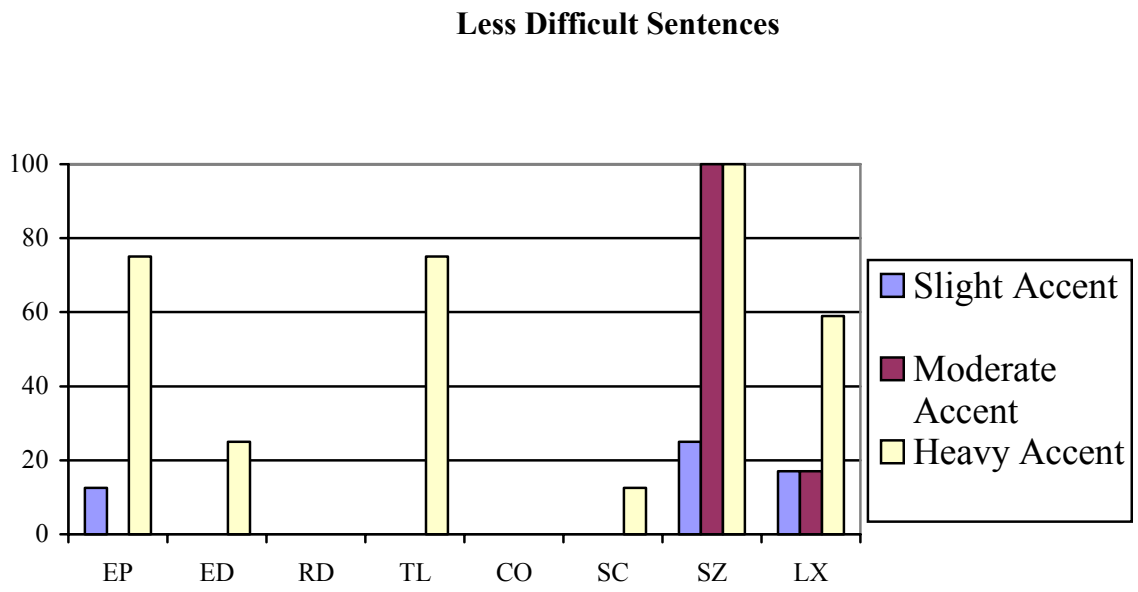
To test the assumption that the speakers produced more errors on the more difficult sentences and that these errors included the targets, the productions of the targets by the individual speakers were investigated. Table 3 and Figure 6 show the distribution and proportion of the targets for the less and more difficult sentences.

Generally the targets that were most in error by all speaker groups were the tense and lax vowel distinctions, the intervocalic voicing distinctions between /s/ and /z/, and lexical stress differences. However, the heavily accented speaker group produced numerous errors on all the targets except for final consonant deletion. It should be noted that the prediction that speakers would consistently make these errors in all contexts was not supported by the data. Most of the time, speakers tended not to commit the expected errors on the targeted English words. Perhaps this explains why there were no significant differences found between the accentedness ratings for the less and more difficult sentences.

Target	# Targets of This Type	% of Total Targets	# in Less Difficult Sentences	% in Total Less Difficult Sentences	# in More Difficult Sentences	% in Total More Difficult Sentences
EP	12	14.81	4	13.79	8	15.38
ED	9	11.11	4	13.79	5	9.62
RD	5	6.17	1	3.45	4	7.69
TL	15	18.52	4	13.79	11	21.15
CO	8	9.88	4	13.79	4	7.69
SC	13	16.05	4	13.79	9	17.31
SZ	8	9.88	2	6.9	6	11.54
LX	11	13.58	6	20.69	5	9.62
Totals	81		29		52	

EP=epenthetic schwa; ED= -ed ending; RD= vowel reduction; TL= tense-lax; CO= final consonant deletion; SC= /ʃ-tʃ /; SZ, /s-z/; LX= lexical stress.

Table 3. Distribution of Targets For Less and More Difficult Sentences



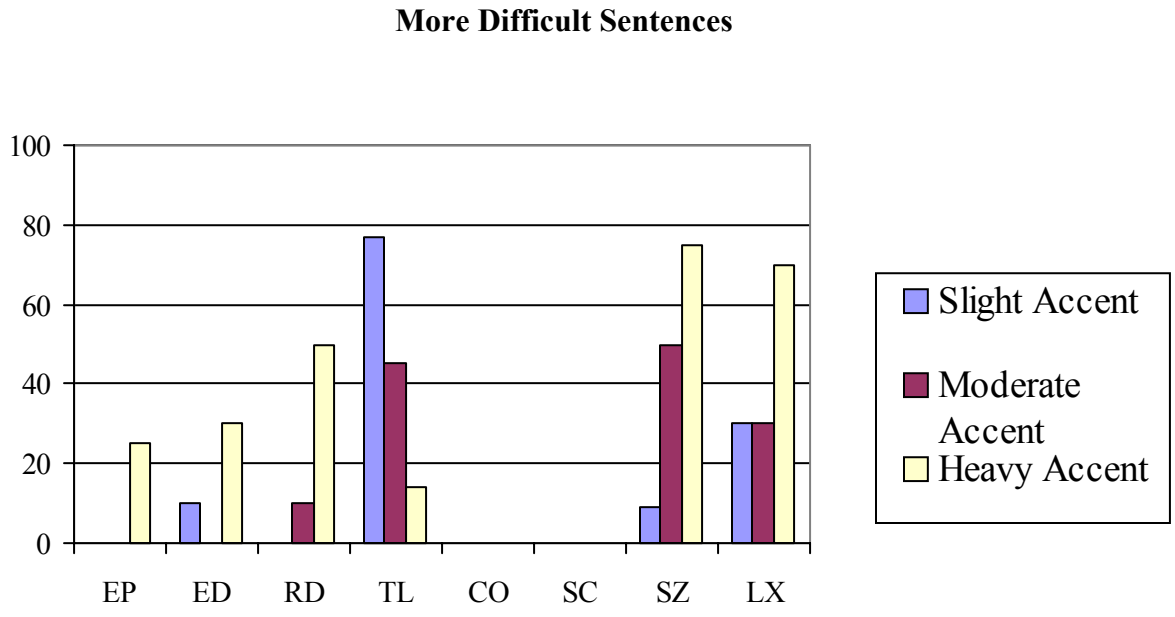


Figure 7. Mean Percent Target Errors for Less and More Difficult Sentences

EP=epenthetic schwa; ED= -ed ending; RD= vowel reduction; TL= tense-lax; CO= final consonant deletion; SC= /ʃ-tʃ /; SZ, /s-z/; LX= lexical stress

It was expected that speakers would insert an epenthetic schwa before consonant clusters in English. The heavily accented speaker group produced this target differently from native norms 75% of the time for the less difficult sentences and 25% of the time for the more difficult sentences. Interestingly, insertion of the epenthetic schwa seemed only to happen when the cluster was preceded by a consonant. For example, in the sentence, *It's not easy to learn Spanish*, an epenthetic schwa was inserted before the word *Spanish* by the heavily accented speakers. However, in the sentence, *She liked the crazy spider*, the target was not inserted since the phoneme /i/ preceded the word *spider*. Perhaps the epenthetic schwa was not inserted before clusters that were preceded by vowels because the vowel served to facilitate the production of the cluster in much the same way as the schwa would have.

The next target investigated was the –ed ending (ED), in which the bilingual speakers were expected to produce /εd/ in regular past tense words instead of /d/ or /t/. For example, the word *liked* was expected to be produced /likεd/ rather than /likt/. The heavily accented speakers produced this target error 25% of the time for the less difficult sentences and 30% of the time for the more difficult sentences. The moderately accented group did not produce the target error at all in either the less or more difficult sentences. The slightly accented speaker group did not produce the target error in the less difficult sentences and only 10% of the time for the more difficult sentences. It was suspected that the orthography may have elicited incorrect production in some of the words. Because the speakers were reading the sentences, they may have pronounced /εd/ in words that were unfamiliar to them, like *stalled.* In contrast, they did not produce the target error when reading more familiar words like *closed* and *called*.

It was also expected that the bilingual speaker groups would produce the vowels in unstressed syllables fully (RD), e.g. *mel[o]ns* for *mel[ə]ns.* However, this error was only produced by the heavily accented speaker group in the more difficult sentences and was not seen in the slightly and moderately accented speaker groups. The heavily accented speaker group also made more errors than the other speaker groups on the other vowel target, the distinction between tense and lax vowel production (TL). It was expected that the bilingual speakers would produce lax English vowels as tense vowels, e.g., *ch[i]p* for *ch[I]p*. The slightly and moderately accented speaker groups only produced this in the more difficult sentences, possibly due to the phonetic context.

It was further expected that the bilingual speakers would drop the word final /s/ or /z/, but none of the bilingual speaker groups produced errors on this target. It was also expected that the speakers would substitute /tʃ/ for the English /ʃ/, but again very few errors were produced on this target. This error only occurred when the target was followed by a tense vowel, e.g. *ship*. It was not seen when the target was followed by a lax vowel, as in *shop*. Interestingly, two of the speakers produced /ʃ/ for /tʃ/ on the words *stopwa<u>tch</u>* and *<u>ch</u>osen*. One of these speakers was from the Dominican Republic and the other was from Mexico City. It is not understood why these speakers made this substitution, especially since their dialects differ and the other speaker from the Dominican Republic did not make this error.

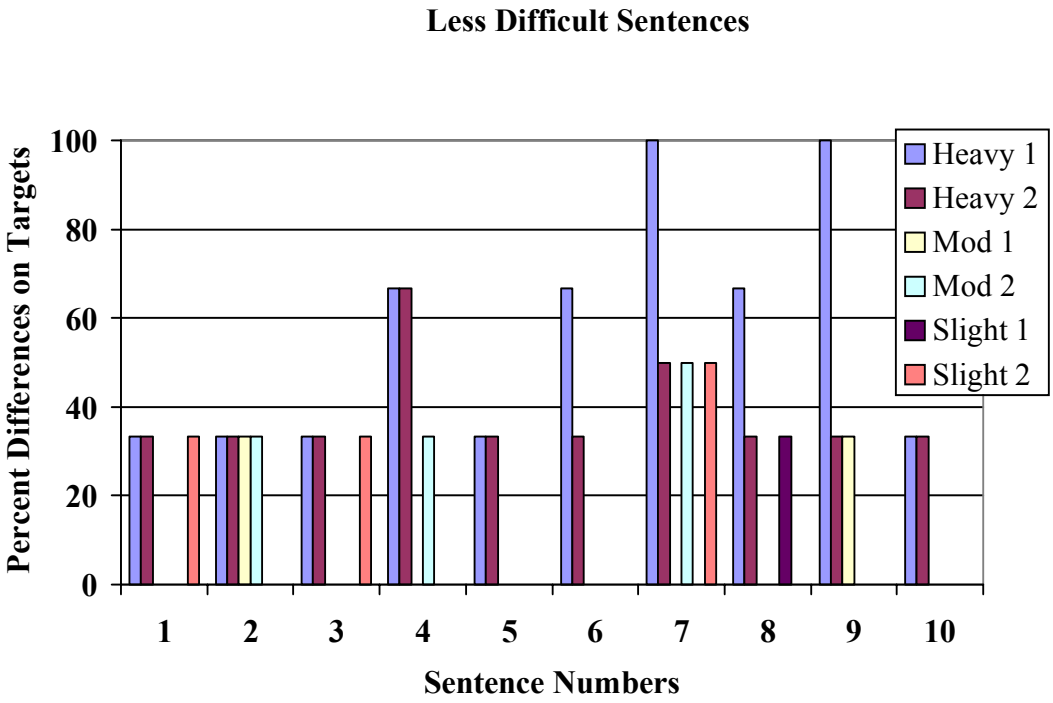
The consonant target that did yield a more significant number of errors was the voicing distinctions between intervocalic /s/ and /z/, e.g., *free[s]er for free[z]er* (SZ). The moderately and heavily accented speaker groups were not able to produce these targets correctly at all in the less difficult sentences. They were more successful in their

productions in the more difficult sentences. The slightly accented speaker group also saw a higher success rate for this target in the more difficult sentences versus the less difficult, possibly due to the phonetic context. Perhaps this helps explain why there were no predicted significant differences found between the less and more difficult sentences; at times, the speakers made more errors on the less difficult sentences. Hence, instead of being influenced by the number of targets present, accentedness may be more word- or context-specific.

Finally, it was expected that the speakers would produce the multi-syllabic English words in the stimuli with different stress patterns (LX) than would native speakers. All speaker groups made errors on this target in both the less and more difficult sentences.

A descriptive exploratory analysis was conducted to determine whether the listeners based their ratings of accentedness on the number of target errors per speakers. To determine this, the total number of target errors per sentence was related to the mean ratings assigned to that sentence by the listener groups. Figure 8 shows the errors per sentence. Where no bar line appears, no errors were made by the speaker for that sentence. The heavily accented speakers made most of the errors in both the less and more difficult sentences. An examination of the ratings assigned to each sentence by the listener groups indicated that the listeners were not rating the accents of the speakers based on the number of target errors produced in each sentence. To illustrate this point, consider the first heavily accented speaker who produced none of the targets correctly in sentence number seven of the less difficult sentences. However, she received ratings comparable to those given to sentence one of the less difficult sentences in which she

produced few errors (see Figure 9). The arbitrary nature of the ratings suggests that the listeners rated the sentences based on something other than the correct or incorrect production of the targets. Perhaps, the listeners rated the sentences based on the speakers' productions that differed from native speaker norms but were not target items not included in the analysis.



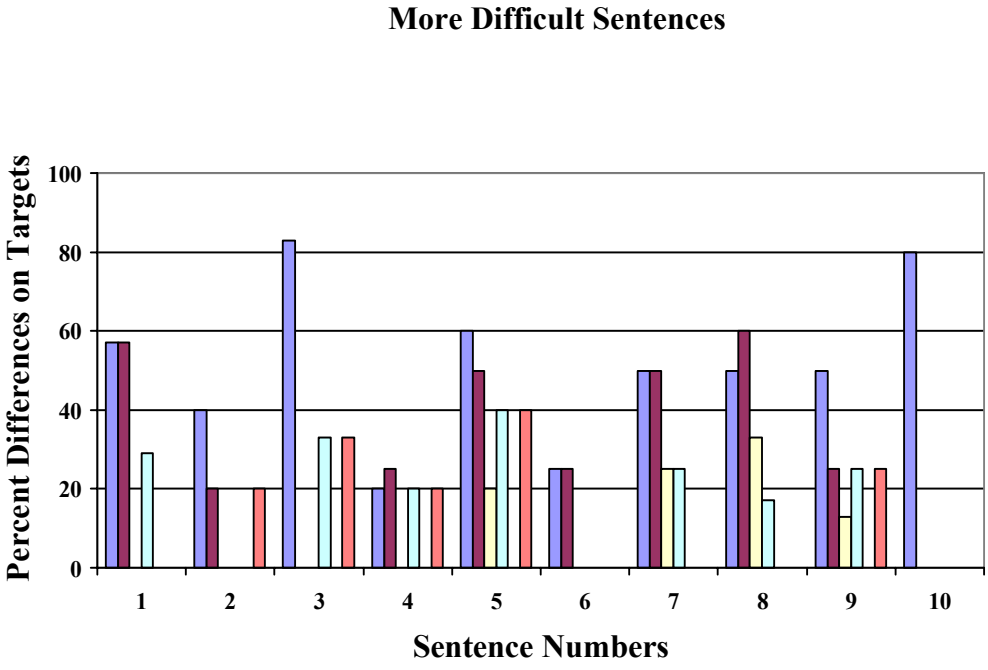


Figure 8. Mean Percent Errors by Individual Speakers for Less and More Difficult Sentences

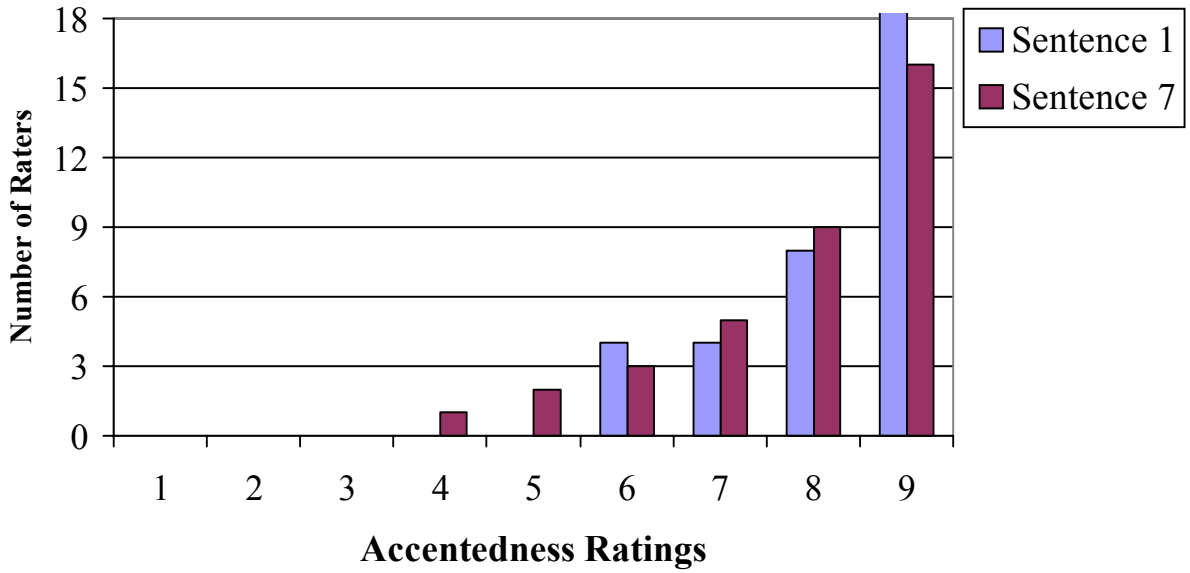


Figure 9. Frequency Distribution of Ratings Given to Less Difficult Sentences 1 and 7 of Heavily Accented Speaker 1.

Summary

Generally, listeners tended to rate the speakers' degree of foreign accent similarly. Amount of daily L1 use was only a significant variable in the ratings of the slightly accented group; here the high L1-use bilingual group gave significantly higher accent ratings than did the native English group, regardless of the level of phonological difficulty. As was evident in the data, listeners were able to differentiate slight, moderate, and heavy accents. Listeners also rated one of the speakers of the heavily accented speaker group as significantly more accented than the other, although this speaker group made roughly the same number of target errors. The level of phonological difficulty in each sentence was not a variable that was generally responsible for differences in accent ratings. The speakers did not make the predicted errors on the

targets on most occasions. In fact, the moderately and slightly accented speaker groups made very few of the predicted errors. Consequently, listeners may have rated the accentedness in the sentences based on something other than the number of target errors contained in each. Generally, all listening groups rated the speaking groups as significantly different from each other, with the least accented speakers receiving the lowest accent ratings and the most heavily accented group receiving the highest.

Chapter 4

Discussion

This study was undertaken to determine whether the degree of perceived foreign accent in an L2 is related to the listener's amount of L1 use and the phonological difficulty of the stimuli. It was hypothesized that individuals who used their L1 less would be better able to gauge the degree of foreign accent in nonnative speakers of their L2 than would individuals who use their L1 more often. This hypothesis was only partially supported. Originally, it was predicted that the high-L1 use listeners would rate the speakers as less accented compared to the ratings given by the low-L1 use listeners. Although the high-L1 use listeners in this study rated the slightly accented speakers as more accented than did the native English listeners, the results can still be interpreted to indicate that the high-L1 use listeners did not possess enough information about how the phonetic segments in English should sound, which was reflected in the differences in their rating patterns compared to the native English listeners. This finding agrees with previous studies showing that as individuals gain experience with the L2, in this study defined as amount of L1 use, they are better able to gauge the degree of foreign accent in the L2 (Flege et al., 1997).

Additionally, it was predicted that the level of phonological difficulty of the stimuli would affect the speakers' production accuracy and would also influence the degree of perceived foreign accent, with more difficult sentences eliciting more production errors and higher accent ratings. Difficulty was determined by how many targets were included in the stimuli. The less difficult sentences were predicted to be perceived as less accented by both bilingual listener groups, with the low-L1 use listener group perceiving them as more accented than the high-L1 use listener group. Finally, it was predicted that the more difficult sentences would be perceived as more accented by all listener groups, specifically, the low L1 use listener group was predicted to give higher accent ratings than the high L1 use group. The hypothesis that the more difficult sentences would receive higher accent ratings was not supported.

Differences Between Listener Groups

The results of this study support Flege's Equivalence Classification hypothesis (Flege et al., 1997), which contends that those who rarely use their first language are better able to gauge non-native accents in their target language than are those who use their first language often. The non-native listeners who spoke their first language less often than their second language became more experienced in English, and they gained more information concerning how the phonetic segments in English should sound. When the accents did not differ significantly from native productions, it became more difficult for the high L1-use listeners to accurately gauge the degree of foreign accent. Flege (1988) hypothesized that native speakers develop detailed phonetic prototypes against which to judge goodness of phones produced by non-native speakers. For non-native speakers, the ability to gauge a foreign accent in English sentences is a skill that develops

slowly with English-language experience. With this experience, adult L2 learners become better able to detect a foreign accent and gauge its strength by establishing these phonetic prototypes.

In this study, listener groups differed significantly in how they rated the slightly accented speaker group. The high-L1-use bilingual group was able to distinguish between all listener groups in the same way as the other listener groups; that is, they gave the slightly accented speaker group the lowest ratings and the heavily accented group the highest ratings. However, the high L1-use bilingual group rated the slightly accented speaker group as significantly more accented than did the native English and low L1-use listener groups. An explanation consistent with both the hypothesis and previous literature suggests that as L2 learners gain experience with the target language, they are better able to gauge accents in that language (Flege, 1984). Although the high L1-use listeners gave higher accent ratings to the slightly accented speaker group, this did not necessarily mean that they were more sensitive to the phonetic and prosodic deviations from the target norms. A possible alternative is that when the speakers' accents differed greatly from the native norms, as was the case for the heavily and moderately accented speaker groups, the high L1-use bilingual group could more easily gauge their degree of foreign accent and thus, performed similarly to the native English and low L1-use bilingual listener groups. However, when the accents differed very little from the native norms, as was the case for the slightly accented speaker group, the high L1-use bilingual listener group was not able to detect the departures from the target language phonetics, and therefore rated the slightly accented speaker group differently from how the native English and low L1-use listener groups rated them.

This makes sense when one recalls the Equivalence Classification hypothesis (Bohn & Flege, 1990; Flege & Eefting, 1987). This hypothesis states that for both perception and production, sounds that are identical in two languages are unlikely to cause a problem, but similar sounds might offer persistent although subtle problems for the second language learner. Similar sounds should be most difficult because they will probably be substituted by a similar sound from the first language, even after the L2 learner has gained considerable experience with the target language. Sounds that are similar between languages are typically vowels, but may also include consonants. Sounds that are completely new, in the sense that they are not equivalent or even similar to sounds from the individual's L1, will be established into a new category as a result of phonetic learning that is not hampered by equivalence classification. As the amount of experience with the L2 increases, individuals will produce second language vowels more like natives.

According to the single system hypothesis, bilinguals have a single phonological system in which the phonetic inventories of both languages reside and they cannot fully isolate either phonetic system (Guion et al., 2000). The hypothesis also asserts that the less L1 there is, the smaller will be its influence on the L2 (Flege et al., 1997). According to Grosjean (1992), the L1 phonetic system influences that of the L2, and the nature of this influence depends on several variables, including the amount and type of use of each language. Generally, the more individuals speak their native language, the stronger will be their accent in their second language (Flege et al., 1997; Guion et al., 2000). Furthermore, this relationship seems to be asymmetrical; although L1 use has an effect on accent in the L2, the L2 typically has a much smaller effect on L1 production. The single

system hypothesis also states that the loss of L1, through high use of the L2, may reduce the degree of perceived foreign accent in the L2.

In terms of the relation between speech perception and speech production, the ability to gauge an accent and discern the phonetic differences between L1 and L2 sounds is determined by the individual's age of learning and the perceived amount of dissimilarity of L2 sounds from the closest L1 sounds (Flege, 1995). The production of an L2 sound will correspond to the mental phonetic category representation that the L2 learner has developed after exposure to the target language. That is why accurate perception of L2 sounds tends to precede their accurate production. However, the only way to really determine whether perception leads production is to develop measures to assess L2 learners' perception and production abilities.

It would therefore be interesting to obtain accent ratings of the bilingual listeners in this study to determine the correlation between their ability to gauge a foreign accent in their L2 and their ability to produce it. A study that tested this idea (Flege, 1988) used bilingual listeners who differed according to the number of years they lived in the United States. The listeners rated the degree of foreign accent in sentences spoken by individuals with varying degrees of accent. The ratings of the bilingual listeners correlated strongly with those given by native English listeners, with the highest correlations between the more experienced listeners and the native English listeners. Following the perceptual tasks, the L2 production of the bilingual listeners was then investigated. It was found that the more and less experienced bilingual listeners had accents that were judged to be equally strong. Thus, the more experienced bilingual listeners were more perceptually sensitive to the phonetic features of English than the less experienced listeners. This may

explain why the listeners in this experiment tended to perform similarly to each other and to the native English listeners in the ability to gauge a foreign accent in speakers with moderate to heavy accents. It was when the speakers' production of the English phonetic features differed only slightly from native norms, as was the case in the slightly accented speaker group, that the less experienced listeners had difficulty gauging the degree of foreign accent. This seems to be true even though the speakers did not make the predicted errors on the targeted features. Clearly, all listeners were sensitive to features of the speakers' productions that signaled a foreign accent.

An alternative explanation considers the reasons that the listeners in this study spoke either Spanish or English most of the time. Perhaps the high L1-use group spoke Spanish most of the time because of their identification with the culture of the L1. Likewise, the low L1-use listeners may want to assimilate more into the culture of the L2, prompting them to use their second language most of the time. The identification with the L1 culture could have resulted in the high L1-use listener group having recognized the slightly accented speakers as Latin, and rated them as accented, more accented than did the other listener groups. In other words, the detection of an accent, albeit slight, was enough to prompt the high L1-use listeners to rate the slightly accented speakers significantly higher than the other listener groups. However, they maintained the relative differences in ratings between slight, moderate, and heavily-accented speakers.

In a tangentially related study, Bresnahan et al. (2002) found that participants with strong ethnic identity deemed American English, which reflected their ethnicity, to be more pleasing and have higher status than the unintelligible foreign accent, whereas those with weak ethnic identity found unintelligible foreign accent more pleasing and attributed

higher status to it than American English. What is relevant to the present study are the authors' conclusions that people with strong ethnic identity may be more attached to their ethnic group and, therefore, more likely to recognize a foreign accent as representing someone in an out-group. Perhaps that is what affected the ratings of the slightly accented speakers given by the high L1-use group in the present study. High L1-use listeners may have spoken Spanish most of the time because of their strong ethnic identity and, therefore, recognized the slightly accented speakers as members of a different ethnic group, as revealed by their dialects, and rated them accordingly. Remember, the speakers and listeners came from various countries of origin with different regional dialects. Perhaps the high L1-use listeners based their ratings on their recognition of the dialects of the speakers, something that would have been difficult for the native English listeners to do.

Another possible explanation is that the ratings of the slightly accented speaker group given by the high-L1 use listener group may have reflected their bias. Remember, listeners were told that all the speakers were non-native speakers of English. Also, the rating scale was labeled "slight Spanish accent" instead of "no Spanish accent". The high-L1 use listener group could have been thought that the slightly accented speakers sounded like native English speakers, but because they were told that all speakers were non-native, they rated them higher than they would otherwise. It would be interesting to see how the listeners would have rated the accents if the background of the speakers was more ambiguous. Likewise, it would have been useful to include a native English speaker group and to modify the rating scale to have the label "no Spanish accent." That

would reveal whether the high-L1 use listeners could discern between native English speakers and those with slight Spanish accents.

Characteristics of the Stimuli

The analysis then turned to the stimuli to determine whether aspects of the speakers' productions affected the listeners' ratings. It was predicted that the more phonologically difficult sentences would result in more mispronunciations, and that the errors made would vary according to the speakers' level of accent, with strongly accented individuals producing more errors on the targets than the slightly accented individuals. However, the results of the analysis revealed that the level of phonological difficulty in each sentence was not generally responsible for differences in accent ratings. The slightly accented speakers produced the highest percentage of correct targets for both the less and more difficult sentences. The moderately accented group had the next highest percentage correct targets, although they differed only somewhat from the slightly accented speaker group. The heavily accented speaker group produced the most target errors for both the less and more difficult sentences. Although the speaker groups performed as expected in terms of which groups committed the most target production errors, the speakers only made the errors some of the time. In fact, the moderately and slightly accented speaker groups made only a small percentage of the predicted errors, and even the heavily accented speaker group made far fewer target errors than expected. Perhaps speakers were being overly careful in their productions because they were reading sentences rather than speaking naturally in conversation. It is also possible that speakers of dialects not used in this study would have produced more target errors than the speakers used here. However, the speakers within each speaker group were from different countries of origin,

and there were no significant within group differences in terms of number of target errors.

Still, listeners were able to gauge the degree of foreign accent in each of the sentences. Consequently, it appears that listeners tended to rate the accentedness in the sentences based on something other than the number of target errors contained in each. Flege (1984) argued that listeners' detection of foreign accent may be based on supra-segmental differences in prosodic features, such as timing, stress, and intonation. Perhaps that is what listeners did here as well. Certainly, they did not rate the degree of foreign accent based on phonemic features targeted alone. For this to have been the case, one would expect the ratings to reflect the speakers' correct productions of the targets, as well as their incorrect productions. However, there seemed to be little relationship between the production accuracy in one sentence and the ratings assigned to it by the groups of listeners. Alternatively, the listeners could have rated the degree of foreign accent based on features and items that were not measured in this study. Perhaps it would have been more useful to classify the sentences as more or less difficult after a transcription was performed. A further explanation is that because the sentence structure varied a lot, some sentences may have been more difficult to process by listeners, so the errors may have been more or less distracting.

This finding supports previous literature that attempted to discern what contributes to the perception of foreign accent (Magen, 1998). In her study on the perception of foreign accented speech, Magen found that listeners rated degree of foreign accent based on factors other than just the targeted phonetic sounds or sequences of sounds. She assessed the contribution of various factors to the perception of foreign

accent by having listeners rate speakers' degree of accent in phrases as originally produced and edited acoustically. The edited versions were intended to more closely resemble the productions of native American English speakers. The edited versions received significantly lower accent ratings for epenthetic schwa, -ed ending, tense-tax distinction, final /s/, /tʃ-ʃ/, and lexical stress for one of the speakers and epenthetic schwa, final /s/, and /tʃ-ʃ/ distinction for the other speaker. Listeners were insensitive to voicing differences between the edited and unedited samples, suggesting that listeners are not as likely to attend to voicing distinctions when they are in the context of larger phonological distinctions. This may help explain the why ratings in this study did not reflect the number of target errors in each sentence. For example, the /s-z/ voicing distinction was the target most often in error for all speakers, yet its mispronunciation did not affect higher accent ratings. It seems possible for this study, as Magen contended for hers, that suprasegmental factors may have contributed to listeners' perceptions of accentedness. In other words, listeners may have been attending to the prosodic features rather than the individual sounds in error. However, as has been noted, the listeners could have been attending to features not measured in this study.

Although when listeners were asked on what they based their ratings, nearly all responded that some of the words sounded "wrong" while others sounded "right." According to the listeners, sentences with words that sounded "wrong" were given higher accent ratings than those with words that sounded "right." However, an investigation into the ratings revealed that they did not always reflect the speakers' production accuracy. Indeed, sentences that had the most sounds in error were not necessarily rated as more accented than sentences that had the fewest target errors. For some targets,

speakers made fewer errors in the more difficult sentences and more errors in the less difficult sentences. Perhaps this helps explain why there were no predicted significant errors found between the less and more difficult sentences. Reasons for this may be the phonetic context of the target within the less versus more difficult sentences. It may be that certain phonetic features are more critical or weighted more heavily. It may also be the case that certain phonetic contexts facilitate correct production of the target and certain phonetic contexts make the target error more likely. For example, speakers tended only to insert the initial epenthetic scwha when the cluster was preceded by a consonant. Certainly, it remains unclear as to what features the listeners attended that influenced their ratings. The listener groups seemed not to base their ratings on the presence of errors on the targeted features. Perhaps they rated degree of accent on phonetic features that were not controlled for, like VOT or vowel duration. An investigation of the phonetic contexts should, therefore, be conducted to determine which hinder or facilitate production.

Future Directions

In this study, there was great variability within the listener groups in terms of how long they had been speaking English and their length of residence. For example, some listeners in the high L1-use listener group (those who spoke Spanish most of the time) reported having learned English as much as 15 years ago. Conversely, some listeners in the low L1-use bilingual group (those who spoke English most of the time) indicated that they had learned English as little as three years ago. This experiment only controlled for the amount of first language use at the time of the experiment, rather than earlier L1 and L2 use or the aggregate amount of L1 and L2 use over the listeners' lifetimes. Therefore,

it is unclear whether the categorization of listeners into high and low L1-use groups is fully accurate. Additionally, the listener groups were formed based on the self-reported amount of L1 use. Although this is consistent with previous literature, self-report may suffer from respondents' miscalculation of their language use and cause listeners to be incorrectly categorized. Considering that this was a perceptual task, it may have been helpful to consider the amount of time listeners spent listening to the L2 also and the type of L2 input they received. Perhaps future studies should include questions on the language background questionnaire regarding amount of time listening to and speaking the L2 since the participant's first exposure. Most ideally, a longitudinal study would elucidate how the nature of L2 input and the amount of L2 listening and speaking affect the ability to gauge foreign accent in the L2.

Considering the results, it may have been helpful to include a native English speaking group. The high L1-use listeners rated the slightly accented speakers significantly higher than did the native English listeners. It would, therefore, have been interesting to compare accent ratings given by the high L1-use listeners to native English speakers versus those given the slightly accented speakers. Such a contrast would reveal whether the reason high L1-use listeners had difficulty gauging the degree of foreign accent in the slightly accented group was that the speakers' productions were so close to native norms.

Finally, it may be helpful to perform acoustic analyses on the speech samples in this study to quantifiably determine where the departures from the native norms occurred. It has been shown that even when listeners do not consciously detect phonetic differences between native and non-native listeners at the segmental level, these differences can still

lead to the detection of accent (Bohn & Flege, 1990). This suggests that listeners assess everything they hear in the speech stimuli, even acoustic aspects that are not overtly evident; this explains the listeners' ability to gauge the degree of foreign accent between speaker groups despite their lack of target error productions. The list of phonemes targeted in this study was not exhaustive; perhaps many other potentially difficult phonemes existed in the sentences.

Conclusions

In summary, the results of this study support the view that as adults use their L2 more often, the better able they will be to perceive and gauge degree of accentedness in the L2. This study was concerned with two major issues: whether ability to gauge the degree of foreign accent in an L2 improves as one uses their L2 more often compared to their L1, and whether the degree of phonological difficulty would affect the production and perception of the stimuli. Effects of L1 use were found for those listeners who spoke Spanish most of the time and, therefore, had the least amount of experience with the L2 in terms of their daily language use. In general, the high L1-use bilingual group rated the slightly accented speakers as more accented than did the native English and the low L1-use bilingual groups. Additionally, the degree of phonological difficulty did not affect the ratings of the sentences. Speakers did not produce the errors as predicted and, subsequently, the listeners did not rate the accentedness in the sentences based on the presence or absence of these errors. The implication is that listeners may be attending to something other than the segmental aspects of the stimuli. Perhaps, as has been suggested in previous literature, listeners do not attend to segmental features, but rather to suprasegmental features (Flege, 1984; Magan, 1998).

The results summarized provide support for the equivalence classification hypothesis. All of the bilingual listeners learned English after the age of twelve. Yet, they performed differently in their ability to gauge the degree of foreign accent in non-native speakers of their L2. Therefore, the notion that a critical period exists for the learning of an L2 can not be accepted in light of the present findings. Many factors in addition to age of learning differentiated the listeners in this study. Those who spoke English most of the time performed more similarly to the native English listeners than did those who spoke Spanish most of the time. The finding suggests that amount of L2 usage has an effect on the ability to perceive L2 phones accurately. This is due perhaps to the refinement of internal phonetic category representations, as described in the equivalence classification hypothesis. However, the amount of daily L1 use was the only experience variable used to categorize the listeners. Certainly, it is possible that prior amounts and types of L2 use, prior and current amounts of L2 listening, nature of L2 input, and social factors, like cultural identification, affected the ability of the listeners to gauge foreign accent. The relative contributions of these factors should be explored and disentangled in future studies to provide a clearer understanding of the processes involved in L2 speech learning.

References

Anderson, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38,* 561-599.

August, D. & Hakuta, K. (1998). Bilingualism and second-language learning. In D. August & K. Hakuta (Eds.), *Educating language-minority children* (pp. 11-19). Washington, D.C.: National University Press.

Bent, T. & Bradlow, A. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114,* 1600-1610.

Best, C., (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore, MD: York Press, Inc.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68*, 706-755.

Bohn, O., & Flege, J. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11,* 303-328.

Bresnahan, M., Ohashi, R., Nebashi, R., Liu, W., & Shearman, S. (2002). Attitudinal and affective response toward accented English. *Language and Communication, 22,* 171-185.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Buffalo, NY: Multilingual Matters.

Flege, J. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76,* 692-707.

Flege, J. (1988). Factors affects degree of perceived accent in English sentences. *Journal of the Acoustical Society of America*, 84, 70-79.

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

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

Flege, J., & Fletcher, K. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America, 91*, 370-389.

Flege, J., Frieda, E., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169-186.

Flege, J., MacKay, I., & Meador, D. (1999a). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America, 106*, 2973-2987.

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

Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34,* 65-89.

Gottfried, T., & Beddor, P. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech, 31*, 57-75.

Griffiths, J. (1967). Rhyming minimal contrast: A simplified diagnostic articulation test. *The Journal of the Acoustical Society of America, 42*, 236-241.

Grosjean, F. (1992). Another view of bilingualism. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 57-63). Amsterdam: North-Holland.

Guion, S., Flege, J., & Loftin, J. (2000). The effect of L1 use on pronunciation in Quichua-Spanishbilinguals. *Journal of Phonetics, 28*, 27-42.

Hill, J. H. (1970). Foreign accents, language acquisition, and cerebral dominance revisited. *Language Learning 20,* 237-248.

Iglesias, A., & Anderson, N. (1993). Dialectical variations. In J. Bernthal and N. Bankson (Eds.), *Articulation and phonological disorders* (pp. 147-161). Englewood Cliffs, NJ: Prentice Hall.

Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturaitonal state on the acquisition of English as a second Language. *Cognitive Psychology, 21*, 60-99.

Johnson, J. & Newpoot, E. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition, 39*, 215-258.

Jones, R. (1997). Beyond "Listen and repeat": Pronunciation teaching materials and theories of second language acquisition. *System, 25,* 103-112.

Krashen, S. D. (1973). Lateralization, language learning and the critical period: Some new evidence. *Language Learning, 23*, 63-74.

Lenneberg, E.H. (1967). *Biological foundations of language.* New York: Wiley.

Levine, Hullutt, (2002).

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251-285.

Magan, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics, 26*, 381-400.

Matsura, H., Chiba, R., & Fujieda, M. (1999). Intelligibility and comprehensibility of American and Irish Englishes in Japan. *World Englishes, 18*, 49-62.

Myers, J., & Well, A. (1995). *Research design and statistical analysis*. Hillsdale, NJ: Lawrence Erlbaum Associated, Inc.

Munro, M., & Derwing, T. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48,* 159-182.

Munro, M., & Derwing, T. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition, 23,* 451-468.

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

Neufeld, G. (1979). Towards a theory of language learning ability. *Language Learning, 29,* 227-241.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30*, 449-472.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191-215.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning, 19*, 245-254.

Snow, C., & Hoefnagle-Hohle, M. (1978). The critical age for second-language acquisition: Evidence from second language learning. *Child Development, 49*, 1114-1128.

Southwood, H., & Flege, J. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics and Phonetics, 13*, 335-349.

Stockwell, R., & Bowen, J. (1965). *The sounds of English and Spanish*. Chicago: The University of Chicago Press.

Tahta, S., Wood, M., & Lowenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to the second language. *Language and Speech, 24,* 265-272.

Wijngaarder, S., Steeneken, H., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of Acoustical Society of America, 111,* 1906-1916.

Zecker, L. (2004). Learning to read and write in two languages: The development of early biliteracy abilities. In C.A. Stone, E.R. Silliman, B. Ehren, & K. Apel, (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 248-266). New York: Guilford Press.

Appendices

Appendix A: Study Advertisement

Earn Extra Credit Points !!

Participate in a study on Foreign Accents

Help us to understand the factors that influence foreign accent.

We need monolingual English speakers and bilingual speakers whose first language is Spanish and second language is English.

If you're interested, please contact Astrid Doty

at azerla@helios.acomp.usf.edu

Appendix B: Practice Sentences

1. The dog bit the man.
2. The boy hits the ball.
3. It is a hot day.
4. The bike is red.
5. I like candy.
6. The cat is playing.
7. I feel sick today.
8. Sometimes we go to the park.
9. My blue car is fast.
10. Please don’t eat the apples

Appendix C: Stimulus Sentences

Less phonologicalally difficult sentences

LX EP CO
The operator stands by. (7)

LX RD CO
The doctor had questions for me. (8)

TL/SC LX
He wished for conversation. (8)

CO SZ EP
It's not easy to learn Spanish. (8)

SC TL CO
He showed it to Elizabeth's mom and dad. (10)

SC EP ED
The shirt is stained with mud. (7)

ED SZ
He stopped freezing it for me. (9)

LX EP ED
Two automobiles stalled. (7)

SC/TL LX
The ship-builder and the sailor work. (9)

LX ED TL
The hospital was called by him. (8)

More phonologicalally difficult sentences

SC LX CO EP TL/SZ ED
The shop-keeper's store is closed. (7)

SC ED LX/SC/TL
She earned one fellowship that year. (8)

TL/SZ/TL/RD/EP ED
Opposition stalled the plan. (7)

EP SC LX CO TL
Stop and Shop supermarkets are big. (9)

EP ED LX TL SC
Steve liked good television shows. (8)

SC ED EP SZ
She passed the stopwatch to the boys. (8)

SC ED SZ EP
She liked the crazy spider. (7)

SZ TL/RD/CO TL SC
Frozen melons are in the shop. (9)

SC RD/TL/LX/TL EP/TL/CO
She was confident in her skills. (8)

SC RD RD SZ EP
She heard the professor chosen to speak. (10)

EP=epenthetic schwa; ED= -ed ending; RD= vowel reduction; TL= tense-lax; CO= final consonant deletion; SC= /ʃ-tʃ /; SZ, /s-z/; LX= lexical stress.

Number in parentheses indicates number of syllables per sentence.

Appendix D: Instructions To Speakers

Explanation of study. We interested in foreign accents and what makes them more or less noticeable.

Language background questionnaire. Let's complete this questionnaire together that tells us a little about yourself and your Spanish and English use. Once I have looked it over, we can continue with the rest of the study.

Recording of sentences. Have a seat in the booth. I am going to read you some instructions. Here is a list of ten practice sentences and the 20 sentences we will use for the study. You are to put on these headphones and then read the practice sentences three times each. After you finish reading the sentence once, take a breath or pause for just a moment and then go on to the next sentence. You will do the same for the 20 sentences that are part of the study. Let's practice with the first five sentences. Please put on your headphones and make sure they are comfortable. Now read each sentence three times in a normal conversational rate and volume. *(Experimenter may demonstrate with the first practice sentence if necessary. After the participant completes the practice trials, ask if she has any questions before proceeding).* Now let's go on to the next set of 20 questions and read them in the same way you did the five practice sentences. Remember, after you finish reading one sentence, please pause and take a breath before reading the next sentence. When you are finished reading all 20 sentences three times each you may leave.

Appendix E: Instructions to listeners

Explanation of study. We are interested in foreign accents and what makes them more or less noticeable.

Language background questionnaire. Let's complete this interview that tells us a little about yourself and your Spanish and English use. Once I have looked it over, we can continue with the rest of the study.

Griffith Listening Test. For the first part of the study, you will see four words displayed on the computer screen. You will hear one of these words spoken. Your job is to select the word you thought you heard by clicking on the word with the mouse. You will only hear the word once. You will do some practice items first. After you completed the practice, we will go on to the real items. When you have finished this part, feel free to take a short break and then return to the computer.

Rating of sentences. You will hear sentences spoken by people who are native Spanish speakers who have learned English as a second language. Your job is to rate each sentence according to how much of a Spanish accent you think each person has. There are no right or wrong answers, just your opinion about how much of an accent you think each person has. You will use this scale on the computer and click with the mouse to the very far left of the screen if you think the person speaks English with very little Spanish accent. You will click to the very far right of the screen if you think the person speaks English with a very heavy or strong Spanish accent. Click anywhere along the scale depending on how strong you think the speaker's accent is. Click "Accept" at the top of the screen after you have rated the sentence. Let's practice first. *(Present the ten practice trials.)* You are to rate each one individually and not base your judgement on the one

before it. Now let us get started on the real sentences. Remember, rate each sentence independently. You may drag the cursor anywhere along the scale to show how much of an accent you think the speaker has. You may advance to the next sentence by clicking on “Accept” at the top of the screen, the computer will not let you advance until it has played the entire sentence.