

7-19-2007

Clear Speech Effects for Vowels Produced by Monolingual and Bilingual Talkers

Teresa DeMasi
University of South Florida

Follow this and additional works at: <https://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#)

Scholar Commons Citation

DeMasi, Teresa, "Clear Speech Effects for Vowels Produced by Monolingual and Bilingual Talkers" (2007). *Graduate Theses and Dissertations*.
<https://scholarcommons.usf.edu/etd/693>

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.

Clear Speech Effects for Vowels Produced by Monolingual and Bilingual Talkers

by

Teresa DeMasi

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

Major Professor: Catherine L. Rogers, Ph.D.
Jean Krause, Ph.D.
Stefan Frisch, Ph.D.

Date of Approval:
July 19, 2007

Keywords: intelligibility, perception, production, second language, communication

© Copyright 2007, Teresa DeMasi

Dedication

I dedicate this thesis to anyone who ever thought they could not achieve the impossible.

If you want something badly enough, you truly can accomplish anything you put your
mind to.

Acknowledgements

The research in this thesis was supported by NIH-NIDCD grant #5R03 DC005561 to Catherine L. Rogers, Ph.D. and by a CAS travel award.

This thesis would not have been possible if it weren't for a number of very special people in my life. First and foremost, I would like to thank Dr. Catherine Rogers. Besides being my advisor and professor, you have also been my mentor and my friend. You've taught me to have confidence in myself and you've made me realize that it is possible to accomplish a difficult task, despite the many challenges that come along the way. I truly admire you for your patience, your understanding, and your dedication to your work. You've inspired me to consider pursuing research in the future. You've always been there for me, especially through hard times and I would just like to say THANK YOU!!!

I would also like to thank my committee members, Dr. Jean Krause and Dr. Stefan Frisch. Thank you for your support in writing this thesis and for taking the time to read my revisions at the last minute!

The many helpful comments and suggestions of Dr. Diane Kewley-Port are gratefully acknowledged, as well.

Next, I would like to thank my lab partners, Merete Møller Glassbrenner and Michelle Bianchi. Merete...you were there from the first day I stepped foot in the lab and you have always been there to address my questions and concerns. Michelle...you've been my partner in crime and you were there when I thought I couldn't make it on my own. Thank you both for being my friends and helping me get through life when being far away from home. I could not have gotten through those endless hours of editing and acoustic analysis without you both...it's over!

Next, I would like to thank my family. Mom and Dad...thank you for your unconditional love and support despite my decision to pursue my graduate degree 1200 miles away from home. Philip...thank you for being there as my little brother. Grandma...thank you for always thinking of me and sending me letters even though I didn't always write back. In case I haven't said this already, I love each and every one of you.

Next, I would like to thank my two best friends. Diana...thank you for always listening to me complain about how much homework I have and for yelling at me when procrastination takes over. Yasmin...I know you've been through difficult times while I was away at school. Thank you for being so strong through it all...I truly admire you for that. Thank you both for being there for me even when I don't return phone calls! Thank you to all of my professors and clinical supervisors at USF. You have provided me with knowledge and skills that I can take with me throughout my future endeavors. Thank you to all of my clients...you are the reason why I chose this profession. Last, but not least, I would like to extend a huge THANK YOU to the class of 2007. It's been a long road, but...WE DID IT!!!

Table of Contents

List of Tables	iii
List of Figures	iv
Abstract	v
Chapter One: Introduction	1
Factors Affecting Intelligibility	1
Talker-Related Factors	1
Listener-Related Factors	2
Environmental Factors	2
Linguistic Factors	3
Two Factors Manipulated in Present Study	4
Bilingualism Factor	4
Three aspects of language	4
Source-filter-theory	5
Phonetic differences between native and non-native speech	6
Vowels	7
Speech learning model	10
Speaking Style Factor	11
Clear speech research	12
Intelligibility in Noise	13
Clear Speech Benefit	14
Purpose of Present Study	15
Research Goals of Present Study	16
Hypotheses of Present Study	17
Chapter Two: Method	18
Participants: Talkers	18
Participants: Listeners	21
Stimuli	25
Conversational List	26
Clear Speech List	26
Recording, Instrumentation, and Procedures	26
Isolation of /bVd/ Words	29
Mixing of Noise	30
Instrumentation for Perception Data Collection	31
Perception Procedures	33

Chapter Three: Results	37
Chapter Four: Discussion	51
Summary of Results	52
Hypotheses	56
Comparison with Previous Research	57
Implications	60
Limitations	61
Future Research	62
References	64
Appendices	68
Appendix A.1: Consent Form for Production Experiment (Monolingual)	69
Appendix A.2: Consent Form for Production Experiment (Bilingual)	72
Appendix B.1: Language Background Questionnaire for Monolinguals	75
Appendix B.2: Language Background Questionnaire for Bilinguals	77
Appendix C: Race/Ethnicity Form	80
Appendix D: Consent Form for Listening Experiment	82
Appendix E: Receipt for Production and Perception Experiments	85
Appendix F: Instructions Handout for Listening Experiment	86

List of Tables

Table 1	Demographic Data for Early Bilingual Talkers	22
Table 2	Demographic Data for Late Bilingual Talkers	24
Table 3	Results of Three-Way ANOVA of the Effects of Talker Group, Speaking Style and Target Word	38

List of Figures

Figure 1.	Layout for the six-alternative forced-choice task completed by all listeners.	33
Figure 2.	Mean percent-correct word-identification scores for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB), for both speaking styles.	39
Figure 3.	Mean percent-correct word-identification scores for “bead” (panel A) and “bid” (panel B) for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB).	42
Figure 4.	Mean percent-correct word-identification scores for “bayed” (panel A) and “bed” (panel B) for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB).	43
Figure 5.	Mean percent-correct word-identification scores for “bad” (panel A) and “bod” (panel B) for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB).	44
Figure 6.	Mean percent-correct word-identification scores of each talker in the monolingual (MO) talker group.	46
Figure 7.	Mean percent-correct word-identification scores of each talker in the early bilingual (EB) talker group.	47
Figure 8.	Mean percent-correct word-identification scores of each talker in the late bilingual (LB) talker group.	48

Clear Speech Effects for Vowels Produced by Monolingual and Bilingual Talkers

Teresa DeMasi

ABSTRACT

‘Clear speech’ is a speaking style that talkers often employ when they know they may have trouble being understood, as when speaking in noise or to a person with a hearing loss. When ‘clear speech’ produced by native talkers is presented in noise to native listeners, it has been shown to be about 10-15 percentage points more intelligible, on average, than normally produced speech. Recent research has shown that bilingual listeners may experience a smaller intelligibility benefit than monolingual listeners from ‘clear speech’ produced by monolingual talkers. The present study compares the ability of monolingual and bilingual talkers to *produce* this clear speech intelligibility benefit.

The present study investigates the hypothesis that bilinguals may produce a smaller intelligibility benefit than monolinguals when asked to speak clearly. Three groups of talkers were recorded: 13 monolingual native English speakers, 22 ‘early’ Spanish-English bilinguals, with an age of onset of learning English (AOL) of 12 or earlier, and 14 later Spanish-English bilinguals, with an AOL of 15 or later. Talkers produced the target words “bead, bid, bayed, bed, bad” and “bod” in both clear and conversational speech styles. Two repetitions of each word were mixed with noise and presented to monolingual English-speaking listeners across two days of testing.

Both monolingual and early bilingual talkers showed a similar degree of clear speech benefit in noise (about 5.5%). Later bilinguals were the least intelligible overall and showed a smaller overall clear speech benefit in noise. Surprisingly, early bilinguals were significantly *more* intelligible than monolinguals in both speaking conditions (by about 6.5%). For the later bilinguals only, performance was significantly *worse* for one target word (“bid”) in the clear speech condition than in the normal speech condition.

These data suggest that later bilinguals, but not early bilinguals, may experience a relative disadvantage when speaking in noise, due to a reduced ability to improve intelligibility by speaking more clearly. Therefore, these persons may benefit from communication strategies or accent reduction programs designed to increase their ability to make themselves understood in difficult speaking environments.

Chapter One

Introduction

The goal of oral communication is for the talker's ideas to be understood by the listener. There are three components to oral communication: 1) the speaker, 2) the message, and 3) the listener. The communication process begins with the talker having an idea. Next, the idea is encoded into a message, in this case, a verbal message. The process is complete when the listener receives the message successfully (Smith, 2007). Sometimes the intent of the message is not understood by the listener, which may result in miscommunication. In order to determine the success of the communication exchange, intelligibility is one measure that can be used. Intelligibility is typically measured as the number of words intended by the talker that are correctly perceived by the listener (Miller, Heise & Lichten, 1951).

Factors Affecting Intelligibility

There are several factors that may affect intelligibility. These factors can be talker-related, listener-related, environmental, or linguistic.

Talker-Related Factors

Talker-related factors can include the age, gender, and physical state of the talker. The talker may be tired and fatigued when communicating the message or the talker may have a speech or language impairment, making it difficult for the listener to accurately understand his/her speech. Another talker-related factor is bilingualism. If the talker is speaking a language that is not his/her native language and if the talker does not have

sufficient linguistic experience in the second language, then a breakdown in communication may occur.

Listener-Related Factors

Listener-related factors that impact intelligibility can include the age, hearing ability, and linguistic experience of the listener. Presbycusis, the decrease in hearing acuity with age (Tye-Murray, 2004), may affect the listener's ability to accurately perceive the message. If the listener has a hearing impairment, he/she may require amplification or other accommodations in order to understand the message. If the talker is using a language that is not the listener's native language, the listener may need additional time or resources to fully process the message.

Environmental Factors

Environmental factors can have a negative impact on communication as well. These factors can include background noise and other distractions, which may make it difficult for the listener to hear the message. Imagine having dinner in a crowded restaurant, where the level of background noise is quite intense. Between the extraneous conversations in the room, the clinking of silver and glassware, and your own conversation, it can be a challenging task for a listener to accurately perceive the message being encoded at the dinner table. Rogers, Dalby and Nishi (2004) found that non-native speech embedded in background noise may create higher task demands for the listener than native speech presented in the same environment. In their study, native, high proficiency non-native late learners of English as a second language (who had a relatively mild but detectable foreign accent), and less proficient non-native late learners of English as a second language (who had a moderate to strong foreign accent) were recorded

reading a list of fifty sentences. These sentences were later combined with noise at three different signal-to-noise ratios. The sentences were then presented in quiet and in noise to native listeners, who were instructed to write down what they heard. Intelligibility was determined based on the number of content words accurately identified in the intended order. Their results showed that speech intelligibility decreased with increasing levels of noise for all talker groups. In the quiet listening condition, the intelligibility of the high-proficiency non-natives was similar to that of the natives; however, in the noisiest condition, the intelligibility of the high-proficiency non-natives was similar to that of the less-proficient non-natives. These results suggest that even fluent bilinguals who retain even a relatively mild foreign accent may be less able to make themselves more intelligible under adverse listening conditions.

Linguistic Factors

In addition to talker, listener, and environmental factors, linguistic factors can affect intelligibility as well. These factors include speaking style, word frequency, and sentence complexity. The speaking style that is employed by the talker may be conversational. Conversational speech tends to flow naturally. Sometimes the talker may use a rapid rate of speech, which can be difficult for the listener to follow. Some talkers may use clear speech in order to make themselves better understood under adverse listening conditions. Specific properties of clear speech will be discussed further below. The talker's choice of words can also affect communication. If the talker selects words that occur frequently in the language, there is a better chance the message will be understood; however, if the talker selects words that occur infrequently in the language, there is a higher chance a communication breakdown may occur. Finally, short, simple

sentences typically are easier to understand, whereas longer, more complex sentences may be more difficult to understand.

Two Factors Manipulated in Present Study

There are many factors that can negatively impact the intelligibility of one's speech. In the present study, two of the factors mentioned above were actively manipulated: the talker-related factor of bilingualism and the linguistic factor of speaking style.

Bilingualism Factor

Bilingualism can be defined as being able to use two or more languages on an everyday basis (Grosjean, 1989). Bilinguals can learn two languages either simultaneously or consecutively. In the present study, all of the bilingual participants except one were consecutive learners. The one subject who was a simultaneous learner reported learning Spanish first and learning English between the ages of two and three years. Since this is a crucial time for language development in children, she can be considered a simultaneous learner. When an individual learns two languages consecutively (e.g., learns a second language after achieving proficiency in the native language), he/she may sometimes speak the second language with a foreign accent. According to Rogers, et al. (2004), a foreign accent is perceived when a non-native speaker produces L2 sounds that differ from native-speaker phonetic norms.

Three aspects of language. Many differences exist between native and non-native speech, which can be easily identified when the three main aspects of language are examined. These three aspects are the content, use, and form of language (Smith, 2007). Language content includes semantics, which refers to the vocabulary and meaning

differences between the two languages. Language use includes pragmatics, which refers to the social aspects of language. Language form includes syntax, phonology, and morphology. Syntax includes word order, grammar, and sentence structure. For example, in English an adjective precedes a noun, while in Spanish the opposite is true (e.g. “white house” in English; “casa blanca” in Spanish). Phonology includes the sounds that make up the language. This would be the individual vowels and consonants. For example, Spanish only has five monophthongal vowel sounds in its phonemic inventory (Dalbor, 1969), while English has approximately 12 (Ladefoged, 1982). Both languages contain diphthongs, as well, but they will not be discussed further because they were not under investigation in the present study. Morphology includes the smaller meaningful units of words. For example, adding *-ing* at the end of a verb will make it present progressive tense in English, while adding *-ndo* at the end of a verb in Spanish has the same effect. Differences between the native and non-native languages can all affect the ways in which the non-native speaker may select, order and pronounce words in a second language and all of these factors may affect intelligibility.

Source-filter-theory. When producing speech, there are many articulators involved, which include the lips, teeth, tongue, jaw, hard palate, velum, and pharynx. The position of the articulators in the oral cavity determines the quality of the speech sound that is produced; therefore, articulatory differences produced by non-native speakers may be perceived as a foreign accent. According to the Source-Filter-Theory, the voice is produced at the vocal folds, which is the source (Kent, Dembowski, & Lass, 1996). The speech signal is then filtered through the oral cavity, which is shaped by the articulators to create individual speech sounds. These sounds are radiated from the

mouth, which results in the output signal of connected speech (Kent, et al., 1996). The resonances in the vocal tract are determined by the positions of the articulators; these resonances determine which source frequencies will be enhanced and which will be dampened. Differences in the position of the articulators between native and non-native speakers will result in differences in vocal tract resonances. Thus, articulatory differences can cause the output speech signal to sound distorted or accented if it does not match the sound pattern that native speakers are used to hearing. As a result, this mismatch of sound patterns may lead to reduced intelligibility of the speech signal.

Phonetic differences between native and non-native speech. Clearly, there are many differences across languages. However, the present research study will only examine intelligibility differences related to pronunciation. There are various types of phonetic differences between native and non-native speech. Non-native talkers may delete non-native sounds, substitute native sounds for non-native sounds, and distort non-native sounds. These differences from native pronunciations can occur amongst consonants or vowels. Prosodic variations are evident across languages, as well. An example of a prosodic difference would be the tonal properties that are present in Mandarin Chinese. Since this is a tone language, two different words may be pronounced the same, but when they are produced with different intonation patterns, the meaning of the word changes. When learning a language such as English, a native Mandarin Chinese speaker may transfer prosodic elements from their native language to English, thereby producing certain English words with incorrect intonation patterns. An example of a consonant difference would be the production of the English glides, /r/ and /l/, by native Japanese speakers. Japanese does have one glide-like phoneme, which is often described

as being between /r/ and /l/ (Small, 1999). Therefore, distinct production of these two English sounds presents difficulty for Japanese speakers (Flege, Bohn, & Jang, 1997). An example of a vowel difference would be native German speakers' difficulty producing the English /ɛ-æ/ contrast, or native Spanish speakers' difficulty producing the English /i-I/ contrast (Flege, et al., 1997). Rogers and Dalby (2005) found that vowel performance on a minimal-pairs probe list produced by Mandarin-accented English (MAE) speakers was a better predictor of connected-speech intelligibility than was consonant performance on the word list. This suggests that vowels may be of particular importance in intelligibility for non-native speakers. As a result of this research, vowel effects were selected to be studied in the present study.

Vowels. The physical characteristics of vowels contribute to their intelligibility. These characteristics include temporal, spectral, and dynamic cues. Based on analyses of American English vowels, conducted by Hillenbrand, Getty, Clark, and Wheeler (1995), these cues can be defined as the following: temporal cues are measured as the duration of the vowel; spectral cues are measured based on a single, steady-state time slice of the vowel; and dynamic formant cues can be measured at 20% and 80% of the vowel duration. If any of these cues are removed or altered, vowels may be less easily identified (Hillenbrand & Nearey, 1999). In the Hillenbrand and Nearey (1999) study, twenty listeners participated in a vowel identification task. Vowels were embedded in /hVd/ words and were presented in three forms: natural vowel, original formant (OF) synthetic vowel, and flat formant (FF) synthetic vowel. The natural vowel stimuli were unaltered. The flat formant synthetic vowel was created using formant measurements of the original vowel. The flat formant vowel was synthesized using fixed formants

measured at the steadiest segment of the original vowel. The original formant synthetic stimuli were created by modeling the full detail of the formant traces measured from the database created by Hillenbrand et al. (1995), updated at 8 ms intervals. Results of the experiment revealed that OF synthetic vowels were identified with greater accuracy than FF vowels; and naturally produced vowels were identified with greater accuracy than OF synthetic vowels. Even FF vowels, however, were identified with relatively high accuracy (about 73.8% on average). These results imply that both spectral and dynamic cues both play an important role in vowel identification.

These effects of these cues were also investigated in a vowel identification task conducted by Ferguson and Kewley-Port (2002), which included young normal-hearing (YNH) and elderly hearing-impaired listeners (EHI). Monosyllabic words that were produced in either a clear or conversational speaking style by a single talker were presented in 12-talker babble, and all listeners were instructed to identify the vowel in each word. All words were produced in /bVd/ context and were isolated from a carrier phrase. The YNH group found vowels to be about 15 percentage points more intelligible in the clear speech condition than in the conversational speech condition. The EHI group did not obtain this same clear speech benefit. It should be kept in mind that the listeners were unamplified and therefore, audibility may have been an issue. Both listener groups were found, however, use all three vowel cues studied (spectral target, dynamic formant movement, and duration) to identify vowels in this task. While previous research has studied speech intelligibility of clear and conversational sentences in noise, Ferguson and Kewley-Port (2002) was the first study to focus specifically on vowel identification in noise in both clear and conversational speech.

Studies of foreign accent also have found that differences in the use of spectral and duration cues in vowel perception and production. For example, Bohn and Flege (1997) examined production and perception of English vowels across three groups: native English speakers and two groups of native German speakers with different amounts of L2 exposure. The English vowel /æ/ was compared to /ɛ/ in order to determine if the non-natives could produce and perceive the difference between these two sounds. This pair of vowels was chosen because /æ/ does not exist in the German language but a vowel similar to /ɛ/ does exist in German. Results of the production experiment found that target /æ/ and /ɛ/ produced by the experienced Germans were very similar to those produced by the native English speakers. However, /æ/ and /ɛ/ productions of the inexperienced Germans differed from the other two groups in two ways: 1) the inexperienced Germans produced /æ/ higher in the vowel space (i.e., closer to /ɛ/) and 2) the inexperienced Germans produced a smaller duration contrast between /æ/ and /ɛ/. Therefore, the inexperienced Germans' production of /æ/ was very similar to their production of /ɛ/. Thus, it can be concluded that the inexperienced Germans produced /æ/ differently from the other two groups in terms of both spectral and durational cues.

Non-native speaker perception may also reflect perceptual cue weighting strategies that are different from those of native speakers. This can be seen in the German subjects mentioned above (Bohn & Flege, 1997). Results of the perception experiment showed that the inexperienced Germans relied heavily on duration cues to distinguish /æ/ from /ɛ/, while they relied less on spectral properties. However, the native English speakers and the experienced Germans relied on spectral properties to identify these vowels, while duration differences had very little influence on their perception. This

suggests that the inexperienced Germans were not attending to the appropriate cues when identifying non-native vowels. The difference from native speaker norms in the use of perceptual cues may have influenced their productions of the English vowels /æ/ and /ɛ/. This perception study suggests that the amount of experience a speaker has with the L2 may influence his/her ability to use native-speaker appropriate perceptual cue weighting strategies to identify vowels in a native-like manner.

In another bilingual study, Kewley-Port, Akahane-Yamada, and Aikawa (1996) conducted an experiment to determine which spectral-temporal properties influenced the intelligibility of Japanese-accented English vowels. Japanese speakers were recorded producing /bVt/ words. These words were presented to native English listeners via a minimal-pairs forced-choice perception task. There were three major results discovered in this experiment. First, Japanese speakers were not able to produce the spectral property of English vowels adequately to differentiate vowels that are located close together in the vowel space. Second, Japanese speakers were sometimes able to produce the dynamic contrasts among English vowels, but not always. Third, Japanese speakers were able to produce the correct durational cues for the English vowels. In order to determine how the acoustic properties contributed to intelligibility, regression analyses were performed. Results of these analyses showed that spectral and dynamic properties of vowels were more important than duration cues in the intelligibility of Japanese-accented English vowels to native English-speaking listeners.

Speech learning model. Inaccurate production of American English (AE) vowels by bilingual talkers may contribute to a foreign accent. As mentioned previously, American English has about 12 monophthongal vowels, whereas Spanish only has five

monophthongal vowels. According to Flege's (1995) Speech Learning Model, as children learn their native language, vowel categories are created in their brains, for each of the vowel sounds that occur in their native language. Based on this theory, native English speakers have at least 12 vowel categories in their vowel space, while native Spanish speakers have only five categories in the corresponding space. This difference in vowel inventory between the two languages suggests that Spanish speakers may have difficulty producing some English vowels accurately. The native vowel categories are considered "old" categories. As bilinguals learn a second language and develop proficiency with their L2, they develop "new" categories, which contain the vowel sounds of the L2 that do not occur in their L1. The more distinction there is between the "old" and "new" categories, the more native-like the second language learner will sound. If an "old" category is heard as similar to a new L2 sound, the old category may be used in its place; if the differences between the non-native's "old" productions and the L2-native pronunciation of the new L2 sound are noticeable to native speakers of the L2, these differences will be perceived as a foreign accent. This increased accentedness may in turn contribute to reduced speech intelligibility.

Speaking Style Factor

While the talker-related factor of bilingualism contributes to reduced intelligibility, the linguistic factor of speaking style also affects intelligibility. Normal everyday speech patterns are known as conversational speech. Another type of speaking style that can be employed by talkers is clear speech. Clear speech is a style of speech that talkers may use when they know they may have trouble being understood such as when talking in a noisy listening environment or when talking to a listener who has a

hearing impairment. Picheny, Durlach, and Braida (1986) conducted acoustic analyses of conversational and clear speech tokens produced by native speakers recorded from a previous study (Picheny, Durlach, & Braida, 1985). Several differences between clear and conversational speech were found. These differences include the following: 1) speaking rate is slower in clear speech; 2) in conversational speech, unstressed vowels are modified (reduced) and final stops are usually not released, while in clear speech, unstressed vowels are not modified as much and more final consonants are released; and 3) the root mean square (RMS) intensities for stop consonants are greater in clear speech than in conversational speech. According to Bradlow and Bent (2002), clear speech consists of the following: slower rate of speech, more frequent pauses between words, greater intensity, more varied pitch range, greater emphasis on the release of plosives, and less vowel reduction.

Clear speech research. Research has shown that the use of clear speech does improve intelligibility of the speech signal. Picheny, et al. (1985) recorded three native English speakers reading 50 nonsense sentences in both clear and conversational speaking styles. These sentences were presented monaurally via headphones to five listeners with sensorineural hearing loss. All listeners were instructed to either write down the sentence that they heard or to repeat the sentence. Results indicated that for native speakers, clear speech is about 17% more intelligible than conversational speech for sentences in quiet. It is important to keep in mind that these results are based on the perception of listeners with hearing impairment, while the present study used listeners with normal hearing. Also in the Pichney, et al. (1985) study, the researchers examined sentences, while the present study investigates vowels.

Previous research has shown that the clear speech benefit is about 8% for vowels alone (Ferguson, 2004). Ferguson (2004) recorded 41 talkers (21 female, 20 male), reading sentences in both clear and conversational speech conditions. From these sentences, 1640 /bVd/ words were extracted and embedded in 12-talker babble for presentation to seven young adult listeners in a vowel identification task. Results indicated that vowels were more intelligible in the clear speech condition than in the conversational speech condition. A gender effect was also found. Females were significantly more intelligible than males in the clear speech condition, while this same effect was not present in the conversational speech condition. The Ferguson (2004) study concluded that the ability to produce a clear speech benefit implies that the talker has linguistic knowledge of the cues that listeners use to identify vowels; it also implies that the talker has the ability to modify those cues in order to make himself more intelligible.

Intelligibility in Noise

Non-native speakers with a foreign accent may not always be well understood under normal listening conditions, and they may be even more difficult to understand under adverse listening conditions. Rogers, et al. (2004) found that even relatively mildly accented speakers may experience disproportionate decreases in intelligibility in noise compared to native speakers. Reduced intelligibility may cause breakdowns in communication, especially in noisy environments. This may lead to heightened frustration levels for both communicative partners. Additionally, this may even cause non-native speakers to withdraw from social interactions, or it may cause native speakers to exclude non-natives from social interactions.

In the Rogers, et al. (2004) study cited earlier, additional results were found related to the issues of bilingualism and clear speech. Native speakers were found to be more intelligible than non-native speakers overall. Interestingly, it was found that highly proficient non-native speakers performed very similarly to the native speakers under quiet listening conditions; however, when even a relatively small amount of noise was added to the speech signal, the high proficient non-natives experienced a greater intelligibility deficit than did the native speakers. This implies even relatively proficient non-native speakers may benefit from instruction in how to make themselves more intelligible in noisy listening conditions.

Clear Speech Benefit

While it has been documented that native speakers are able to produce a clear speech benefit (Picheny, et al. 1985), it is unknown whether non-native speakers are able to produce this same benefit or the extent to which it may differ amongst bilinguals with varying degrees of L2 proficiency. Bradlow and Bent (2002) suggested that clear speech is native-listener oriented, and will only benefit listeners who are experienced with the sound structures of the target language. In their study, two native English speakers were recorded reading 64 simple sentences in two speech conditions, clear and conversational. These sentences were originally designed for children and, thus, contained a limited vocabulary. The sentences were later embedded in white noise and presented to 32 native and 32 non-native listeners in a sentence recognition task. All non-native listeners came from various language backgrounds and had limited exposure in an English-speaking environment. Results indicated that non-native listeners received a much smaller clear speech benefit in noise than did native listeners. It is important to keep in

mind that the non-native listeners in the Bradlow and Bent (2002) study were not separated by age of acquisition of English. Since early learners tend to be much more native-like than late-learners (Rogers, Lister, Febo, Besing & Abrams, 2006), the question of whether a clear speech benefit differs amongst bilingual listeners from various L2 proficiency levels was not addressed by Bradlow and Bent (2002). Additionally, there has not been any research found to date that has compared native and non-native talkers' ability to produce a clear speech benefit in noise. Previous research has used native talkers. We already know that some non-native listeners receive less of a clear speech benefit in noise than native listeners for some speech materials, but we do not yet know whether non-native talkers are able to produce a clear speech benefit in noise when asked to speak clearly.

Purpose of Present Study

Understanding the degree to which bilinguals can improve intelligibility when asked to speak clearly, compared to monolinguals, can be useful in accent modification therapy for those who seek it. For example, if a speaker is not able to produce and perceive L2 sounds accurately, then there is a possibility that the speaker will not be able to improve clarity when he/she needs to do so. The results of the present study can help to better understand the effect that the instruction to speak clearly will have on the intelligibility of non-native speech.

The present study will also add to our understanding of the problem because a reduced clear speech benefit for non-natives talkers may suggest that bilinguals lack the ability to enhance the appropriate cues when asked to speak more clearly, perhaps reflecting differences in linguistic knowledge. If this is the case, acoustic analyses might

help to determine which cues native and non-native speakers enhance or do not enhance when asked to speak more clearly.

The present study compared the intelligibility of six American English vowels produced by three groups of speakers: monolingual native speakers of English, early Spanish-English bilinguals, and late Spanish-English bilinguals. Spanish-English bilinguals were chosen because they constitute a large portion of the bilingual population in Tampa, Florida, which is where the study was conducted. According to the United States Census Bureau (2000), there are 28 million Spanish-speaking bilinguals in the United States. This rapid growth of Spanish-English bilinguals has given rise to the need for research in the area of speech production in this population. Six vowels were targeted in the present study. They are /i, I, e, ε, æ/ and /a/. These specific vowels were selected because in American English, they represent a variety of high, mid, and low vowels, and a good combination of tense and lax vowels, as well. Also, all of these vowels occur frequently in the English language. Three of these vowels occur in the Spanish language (/i, e, a/) and are therefore often considered “old” vowels in the bilinguals’ L2 vowel repertoire (Flege, 1995). Three of these vowels do not occur in the Spanish language (/I, ε, æ/) and are therefore often considered “new” vowels in the bilinguals’ L2 vowel repertoire (Flege, 1995).

Research Goals of Present Study

Results of this study provide insight on how native Spanish speakers’ productions of AE vowels are perceived by native English listeners and what the differences are in intelligibility across talker groups, speaking styles, and individual vowels. Additionally, results of this study provide information as to how vowels, specifically, may contribute to

non-native talker intelligibility. The following three research goals were addressed in this experiment. First, a comparison across all three talker groups was made to determine the differences in native and non-native talkers' ability to produce a clear speech benefit in noise. Second, a comparison was made of the differences in the effects of clear speech across individual target vowels for each talker group. Third, the performance of individual talkers was compared within each group.

Hypotheses of Present Study

The investigators formulated hypotheses based on their expected findings for this study. First, it was hypothesized that there would be no change in vowel production across conversational and clear speech conditions for the late bilinguals for some target vowels. Second, it was postulated that the early bilinguals would produce a clear speech benefit in noise. Lastly, it was hypothesized that the native speakers would also produce a clear speech benefit in noise.

Chapter Two

Method

Participants: Talkers

The population that was targeted in the first part of this experiment included native monolingual English talkers, highly-experienced early Spanish-English bilinguals, and less-experienced late Spanish-English bilinguals. Participants were recruited through campus advertisements including flyers, the student newspaper, and e-mail. Groups were selected based on language experience. Native English-speaking monolinguals were placed in the monolingual group (MO); native Spanish talkers who began learning English intensively prior to the age of 12 years were placed in the early bilingual (EB) group; and native Spanish talkers who began learning English intensively after age 15 years were placed in the late bilingual (LB) group. Additionally, in order to qualify for the EB group, participants had to rate themselves as English dominant or balanced in at least two of four modalities (listening, speaking, reading, or writing) and one of these modalities had to be non-print (listening or speaking). The ages of 12-15 years were chosen as a separation range between the two bilingual groups because previous research has found that most people who acquire a second language after age 15 years will speak the L2 with a perceptible foreign accent (Flege, et al., 1997). In addition, those who arrive in the U.S. before age 12 will typically have at least 6 years of schooling in English and will therefore have a relatively long period of immersion by age 18.

The inclusion/exclusion criteria for all subjects included the following: either gender could participate; no subject with a prior history of speech or hearing impairment was included; all subjects had to fall within a certain age range (native talkers had to be between the ages of 18 and 45 years, while bilingual talkers had to be between the ages of 18 and 60 years); and all subjects had to pass a pure-tone hearing screening at 25dB HL at 500, 1000, 2000, and 4000 Hz. The audiometer was calibrated to ANSI 1989 standards. Additionally, it was preferred that all native talkers be from the Tampa Bay area or have lived in the state of Florida for at least five years. This last criterion was chosen in order for the native talker group to have a relatively consistent and standard regional accent. Some native talkers were allowed to participate even if they did not meet this criterion, as long as they did not possess a strong regional accent. Given the demographics of Spanish talkers in the Tampa Bay area, it was not practical to find a large subset of native Spanish talker subjects who spoke the same dialect; therefore, all New World varieties of Spanish were deemed acceptable for participation in this experiment.

For the native English talkers, no native/fluent talkers of any language other than English were allowed to participate; they could not have exposure to a foreign language other than studying a foreign language in high school or college to meet graduation requirements. Additional criteria for the bilingual talkers included the following: Spanish had to be their native language; they could not speak the Castilian dialect which is spoken in Spain, since this is very different from Spanish dialects of the Americas; they should have learned English either as a second language or simultaneously with Spanish; they could not speak any languages other than English and Spanish; they could not have

exposure to any other language other than a foreign language studied in high school or college to meet graduation requirements. All talkers were paid \$10 per hour for their time, with the exception of one early bilingual participant who opted to volunteer her time for this experiment. A receipt was provided to those participants who received monetary compensation. This receipt can be found in Appendix E.

Original participants included 24 monolinguals (3 male, 21 female) with an age range of 18-38 years (mean 22.6 years); 33 early bilinguals (6 male, 27 female) with an age range of 18-35 years (mean 21.5 years); and 21 late bilinguals (5 male, 16 female) with an age range of 19-57 years (mean 29.4 years). Some subjects were dropped for various reasons. Although both genders were eligible to participate, all male subjects were eventually dropped due to a disproportionate number of male volunteers. In previous research, Ferguson (2004) found a gender effect on intelligibility between clear and conversational speech. In order not to skew our data, we decided it was best to drop all male participants. Other participants were dropped if they failed to meet the inclusion and exclusion criteria listed above. In some instances, participants seemed to fit all criteria, but later on, after careful review of their language background questionnaires, it was clear that some of them, in fact, did not meet all criteria for inclusion, and these participants were dropped. Of the 24 monolingual participants, 11 of them were dropped. Of the 33 early bilingual participants, 11 were dropped. Of the 21 late bilingual participants, seven were dropped. Participants were dropped for the following reasons: including the following: prior history of speech or hearing impairment; poor vocal quality (e.g., hoarseness) as observed by investigators; presence of a regional dialect of American English or noticeable pronunciation patterns of African American Vernacular

English (AAVE), as subjectively judged by investigators; unclear age of acquisition of English, as reported on language background questionnaire. Some bilingual participants were also dropped because they did not meet criteria for placement in either bilingual group, as per their report on language background questionnaire (e.g. for inclusion in EB group, participants must report being English dominant for at least two of four modalities: speaking, listening, reading, and writing; one of which must be either speaking or listening).

After subjects were dropped, there were 49 participants remaining. These 49 participants included 13 native English talkers with an age range of 18-34 years (mean 22.38 years), 22 early bilinguals with an age range of 18-35 years (mean 21.73 years), and 14 late bilinguals with an age range of 19-57 years (mean 29.57 years). Additional demographic information for the two groups of bilingual talkers is provided in Tables 1 and 2 on the following pages.

Participants: Listeners

The population that was targeted in the second part of this experiment included monolingual, native English-speaking listeners. Inclusion/exclusion criteria included the following: either gender could participate; no subject with a history of speech or hearing impairment could participate; all participants had to be between the ages of 18 and 45 years; participants could not speak any language other than English fluently; they could not have a strong regional accent; they could not rate themselves as having less than native-like proficiency in English; and they had to pass a pure-tone hearing screening presented binaurally at 25dB at 500, 1000, 2000, and 4000 Hz.

Table 1

Demographic Data for Early Bilingual Talkers

Code	Age	Born/ Raised in US?	Country	AOI	Language background information			
					Language most comfortable for:			
					Speak	Listen	Read	Write
EB02	18	N	Dom.Rep.	7	E	E	E	E
EB04	18	N	Mexico	6	E	E	E	E
EB05	19	Y	Cuba	4.5	E	E	E	E
EB06	19	N	Mexico	5	B	B	E	E
EB08	19	N	Nicaragua	8	E	E	E	E
EB10	19	Y	Nicaragua	6	B	B	B	B
EB11	20	Y	Cuba	6	E	E	E	E
EB12	24	N	Puerto Rico	10	E	E	E	E
EB14	20	Y	Cuba	4	E	E	E	E
EB16	19	Y	Mexico	6	S	E	E	E
EB17	19	Y	Cuba	4	E	E	E	E
EB18	35	N	Venezuela	9	E	S	S	E
EB19	18	Y	Cuba	4	E	E	E	E
EB20	27	N	Venezuela	4	B	E	E	B
EB23	29	N	Puerto Rico	9	E	E	E	E
EB24	26	Y	Colombia	5	E	E	E	E
EB25	21	N	Colombia	11	E	E	E	E

EB26	26	N	Venezuela	12	B	B	E	E
EB29	19	Y	Cuba	2	B	B	E	E
EB30	19	N	Venezuela	8	B	B	B	E
EB31	22	Y	Mexico	6	E	E	E	E
EB33	22	N	Colombia	6	S	E	E	S
Avg./	21.7	10 Y;12	3 Colom.4	6.5	14 E;	16 E; 5 B;	19 E; 2	19 E;
Sum.		N	Venez.		6 B; 2	1 S	B; 1 S	2 B; 1
			6 Cuba		S			S
			4 Mexico					
			5 Other					

Note. The columns in the above table contain the subject code (EB: early bilingual), age of subject in years, whether or not the subject was born in the U.S. (Y: yes, N: no), and the country where either the subject or his/her family originates from. AOI indicates the age of onset of immersion in an English-speaking environment. The last four columns indicate the language that the subject reported being more comfortable in for speaking, listening, reading, and writing (E: English, S: Spanish, B: both languages). These data were retrieved from the language background questionnaire completed by all bilingual subjects.

Table 2

Demographic Data for Late Bilingual Talkers

Subject Code	Age	Born/ Raised in US?	Country	AOI	Language background information			
					Language most comfortable for:			
					Speak	Listen	Read	Write
LB01	30	N	Panama	21	E	S	B	B
LB03	19	N	Colombia	15	S	S	S	S
LB06	19	N	Colombia	16	S	S	S	S
LB07	50	N	Colombia	45	S	S	S	S
LB09	21	N	Colombia	20	S	S	E	S
LB10	28	N	Colombia	28	S	S	S	S
LB11	22	N	Colombia	22	S	S	S	S
LB12	35	N	Colombia	35	S	S	S	S
LB13	19	N	Puerto Rico	16	S	S	S	S
LB15	22	N	Colombia	18	S	S	S	S
LB16	49	N	Colombia	46	S	S	S	S
LB18	57	N	Colombia	30	S	S	S	S
LB19	22	N	Cuba	19	S	S	E	E
LB21	21	N	Colombia	18	S	S	S	S
Avg./	29.6	14 N	11 Colom.	24.9	13 S;	14 S	11 S;	12 S;
Sum			1 Cuba		1 E		2 E;	1 E;
			3 Other				1 B	1 B

Note. The columns in the above table contain the subject code (LB: late bilingual), age of subject in years, whether or not the subject was born in the U.S. (Y: yes, N: no), and the country where the subject originates from. AOI indicates the age of onset of immersion in an English-speaking environment. The last four columns indicate the language that the subject reported being more comfortable in for speaking, listening, reading, and writing (E: English, S: Spanish, B: both languages). These data were retrieved from the language background questionnaire completed by all bilingual subjects.

Listeners were recruited through campus advertisements and through the Communication Sciences and Disorders (CSD) department at the University of South Florida. All subjects were given the option of either receiving monetary compensation for their time or receiving extra credit points for class. Those students who chose to receive monetary compensation were paid \$10 per hour for their time, with a \$10 bonus at the completion of the experiment, provided that they showed up to all sessions on time. A receipt was provided to those participants who received monetary compensation. This receipt can be found in Appendix E. Twenty listeners participated (3 male, 17 female) with an age range of 18-33 years (mean 22.45 years). Range of years living in the state of Florida was reported by listeners to be between six months and 23 years.

Stimuli

Each participant from the three talker groups was recorded reading a series of carrier phrases, each of which contained a target /bVd/ word. The talkers were instructed to read the carrier phrases under two conditions: conversational speech and clear speech. Intelligibility was measured by presenting the recorded stimuli to the monolingual English-speaking listeners via a forced-choice perception task. The percent correct on the

forced-choice perception task was used to determine the intelligibility of each of the three groups of talkers, as well as the differences in intelligibility, if any, across the two speech conditions.

Conversational List

For the production part of the experiment, all three groups of talkers were presented with two lists of stimuli. The first list, conversational, contained 84 sentences. Forty-two of these sentences contained a carrier phrase with a /bVd/ word embedded in the middle (e.g. “Say /bid/ again.”). Each /bVd/ word contained one of the following six vowels: /i, I, e, ε, æ/ and /a/. Seven repetitions of each carrier phrase were recorded, for a total of 42 /bVd/ words recorded. The other 42 sentences contained foil words embedded in the same carrier phrase (e.g. “Say cat again.”).

Clear Speech List

The second list, clear speech, contained a total of 42 sentences, each containing a target /bVd/ word embedded in the same carrier phrase that was used in the conversational list. There were also seven repetitions of each of the six vowels for a total of 42 words. Foil words were not used in this list because the investigator did not want to distract the talker from the target form of the /bVd/ word; instead, the investigator wanted the talker to focus on the /bVd/ word for this particular task.

Recording, Instrumentation, and Procedures

All talkers were recorded while seated in a sound-attenuating booth with incandescent lighting. This type of lighting was used to reduce noise from the fan attached to the ceiling light during recording. All stimulus sentences were typed into a PowerPoint presentation file, and they were visually presented to each talker via a flat

panel computer display. An Audio Technica microphone (AT4033) was placed at a 45 degree angle, approximately six inches from the corner of the talker's mouth. All sentences were recorded onto a DSW digital recorder (Roland, VS880 Ex). The signal was recorded between -4 and -12dB on the VU meter of the DSW. A pre-amplifier was used, which supplied phantom power to the microphone (48V), and included a gain control for the microphone.

Prior to recording, all three groups of talkers were presented with a list of twelve words on a piece of paper, which were modeled by the investigator. This list contained the six different /bVd/ words, as well as the six foil words. The target words were typed as "bead, bid, bayed, bed, bad," and "bod." The foil words included: "cat, cut, cap, cup, cape," and "coop." These words were selected because they were different enough from the target /bVd/ form to distract the talker from the target form, but they were similar enough in that they were monosyllabic CVC words. Voiceless stops were chosen for the first and last consonant sounds in order to increase variability of the types of foil words. The vowel centers were chosen by simply picking vowels from various areas of the vowel quadrilateral in order to create a real word. Some of the foils contained vowels that were not investigated in the present study. The idea behind the foils was to distract the talkers away from the target word enough so that they did not pick up on the target word. If the foils were too different from the targets, then the talkers would have been more aware of those differences, and thus, the target words would have been more salient. Each subject was asked to repeat each word after hearing the investigator's pronunciation of each word. This was done to reduce orthographically-related errors in pronunciation.

A practice recording was made first, which contained twelve sentences. Each sentence was presented visually via a flat-panel computer display and aurally over a speaker. The visual and aural stimuli were presented simultaneously. The talker was instructed to repeat each sentence after seeing and hearing it. This was done to familiarize each subject with the task.

After completing the twelve practice items, each talker was instructed to read the sentences in the conversational list in a normal speaking voice. All sentences were presented one at a time and only the printed version of the sentences was presented on these trials. All talkers were given a bottle of water, and were told that they would be reading 84 sentences, with a short water break after every 21 sentences. After completing recording for the conversational list, all talkers were given new instructions for the clear speech list. They were told to speak as clearly as possible, as if talking to someone who was hard-of-hearing or as if someone was having trouble understanding them. This was how the two speech conditions were differentiated. If any of the talkers produced any of the sentences with obvious dysfluencies, they were instructed by the investigator to repeat the sentence.

All recording procedures took approximately one hour per subject and were completed in one session. First, each talker was given paperwork to fill out, including a consent form (see Appendices A.1 and A.2), a language background questionnaire (see Appendices B.1 and B.2), and a race/ethnicity questionnaire (see Appendix C). All paperwork was approved by the Institutional Review Board (IRB). After the talker finished filling out the paperwork, the investigator conducted a pure-tone hearing screening on each subject. All subjects had to pass a pure-tone hearing screening at 25dB

HL at 500, 1000, 2000, and 4000Hz. Next, the word list was read aloud by the investigator and repeated by the subject to reduce orthographical errors. Then, the practice list of twelve sentences was recorded, followed by the conversational list, and then the clear speech list. Water breaks were provided after every 21 sentences for both speech conditions. All recordings were stored on the hard drive of the DSW, and were later transferred to the hard drive of a personal computer. A back-up copy of each recording was saved onto a compact disc.

Isolation of /bVd/ Words

After all recordings were made, the /bVd/ words were isolated in order to be presented to the monolingual native English-speaking listeners for the perception part of the experiment. All recorded /bVd/ words were isolated from their carrier phrase using the Cool Edit 2000 (2000) software program. In order to do this, first, the waveform of the carrier phrase was opened in Cool Edit 2000 (2000) on a computer screen. After playing the phrase and listening to make sure the correct phrase was selected, the beginning of the phrase was deleted by highlighting it with the mouse and pressing the “delete” button on the computer keyboard (e.g. the word “say” in the phrase “Say bead again”). The same was done for the ending of the phrase (e.g. the word “again” in the phrase “Say bead again”). After each word was extracted from its carrier phrase, the entire word was highlighted and the RMS amplitude of the entire remaining file was equalized to 25 dB less than the maximum amplitude so all words would have a constant average intensity. Thus, the /b/ release and up to 10ms of pre-voicing were preserved and the /d/ release and 10ms following the burst were preserved for all talkers.

This procedure was followed for the first five repetitions of all recorded /bVd/ words in both lists. Seven repetitions were recorded in case a talker had any dysfluencies while pronouncing each carrier phrase and because future research on this database may examine trial-to-trial variability in talkers' productions, but only two high-quality exemplars were used for the present study. These exemplars were usually the first two repetitions of each target word, unless there was a problem such as a dysfluency or the presence of extraneous background noise. In total, the stimuli included two repetitions of each of the six target words, for each of the 49 talkers, in each of the two speaking style conditions, which led to 1176 stimuli all together.

Mixing of Noise

The selected target words were mixed with multi-talker babble using a customized MATLAB script. The reason for adding noise to the signal was because ceiling effects would be expected in quiet conditions for both conversational and clear speech tokens for the monolingual and early bilingual talkers (cf. Hillenbrand & Nearey, 1999), perhaps masking improvements in clear speech performance. Based on pilot testing, two signal-to-noise ratios (SNRs) were selected for the listening experiment. For the monolinguals and early bilinguals, a signal-to-noise ratio of -8dB was chosen; whereas, for the late bilinguals, -4dB was chosen. The reason for this difference was to avoid ceiling and floor effects across all groups. In pilot testing on conversational style tokens, the differences in intelligibility across talker groups was found to be so great that no single signal-to-noise ratio was found that avoided both ceiling effects for the monolingual and early bilingual talkers and floor effects for the late bilingual talkers.

The SNRs chosen resulted in approximately equal performance for all three groups in the pilot testing.

For the mixing of the words with noise, two minutes of multi-talker babble from the SPIN sentences (Bilger, Neutzel, Rabinowitz, & Rzeczkowski, 1984) was recorded from a compact disk, amplitude equalized and saved to file. The customized MatLab program randomly selected a section of this two minute babble file that was equal in duration to the duration of the target word file plus 1000 ms, with a new randomly selected segment for each target word. The program then compared the RMS amplitude of the noise section and the word file, scaled the noise to obtain the desired signal-to-noise ratio, mixed the word and noise files so that the word was centered in the noise (with 500 ms preceding and following the word), and rescaled the combined file to the original RMS of the word file (approximately -25 dB from maximum).

A separate randomized mixing of words with noise was completed for each listener, so that the mixing of particularly high peak or low valley in the noise with the vowel of a particular word would not be repeated for every listener. The 1176 resulting stimuli were then saved in a separate directory for each listener on the computers used for presentation of stimuli to the listeners.

Instrumentation for Perception Data Collection

For the perception part of the experiment, a personal computer containing the Praat (Boersma & Weenik, 2006) software program was used. The Praat (Boersma & Weenik, 2006) experiment presentation program, MFC, was used to control presentation of the stimuli.

On presentation, the stimuli were routed to TDT PA5 (TDT System III, 2001) attenuators that were set to the desired attenuation level, whose output was then routed to TDT HB7 headphone buffers, to which the headphones were attached for each station. All stimuli were presented at approximately 70 dB SPL. Equipment was calibrated as follows: a 1000 Hz calibration tone was played; this tone was saved to the same average RMS level as all the word stimuli; and the tone was then played over the right and left headphones in turn in each listening station to a Bruel & Kager model 2235 sound level meter. The reading in dB SPL of each ear was taken for each set of headphones; the reading for both ears combined was averaged (left and right always agreed to within about 2 dB); and the attenuation level needed to get the output to 70 dB SPL was computed. This attenuation level was set on the programmable attenuators and the calibration tone was replayed to the sound level meter to double check that the output was indeed 70 dB SPL. Thus, the average presentation level of the word stimuli to the listeners was equal to that of the calibration tone (70 dB SPL).

All stimuli were output to the listeners via Sennheiser HD265 headphones in a randomized order, with one half of the stimuli (588 trials) presented on each of the two days of listening. The first repetition of each target word was used for one half of the talkers and the second repetition of the target words was used for the other half of the talkers on the first day of testing. The remaining stimuli were presented on the second day of testing. All listeners were seated in a quiet sound-treated room. Each listener was seated at a separate carrel. All listeners were seated at least one carrel apart, and dividers separated each carrel. Each listener's carrel contained a flat screen monitor, keyboard, mouse and headphones.

Perception Procedures

For the perception part of the experiment, monolingual native English-speaking listeners completed a six-alternative forced-choice perception task. The instructions given were: “You are going to hear a word over headphones. You have six choices on the computer screen in front of you. Use the mouse to click on the word you heard.” The six choices were: “bead, bid, bayed, bed, bad,” and “bod.” All listeners performed this task in the same setting. Figure 1 shows the layout of the computer screen for this task.

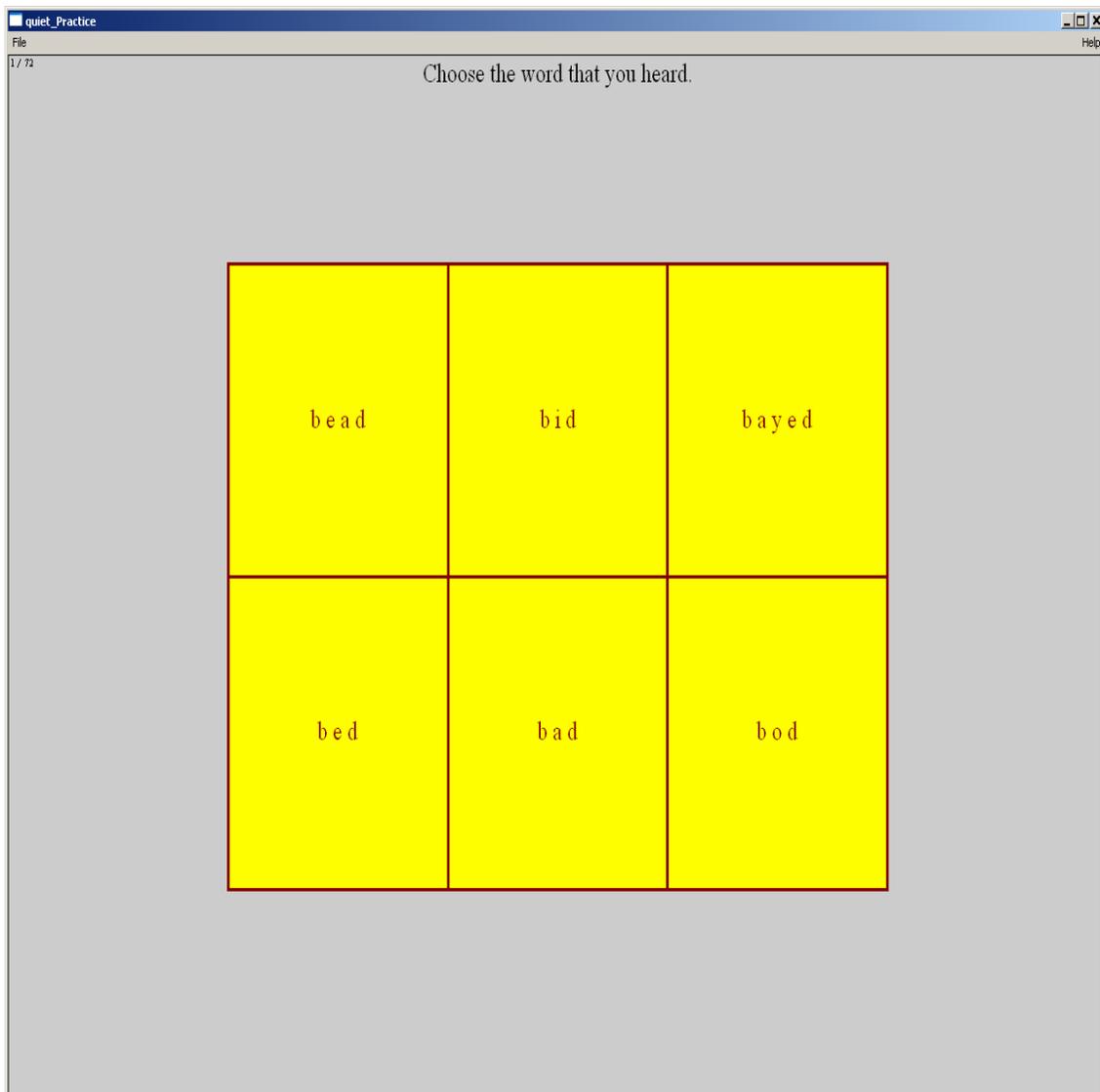


Figure 1. Layout for the six-alternative forced-choice task completed by all listeners.

Due to the number of stimuli presented to each listener, the experiment took place over two separate sessions in order to reduce fatigue. Stimuli were also presented in quiet on two additional days of testing, but these data will not be presented in this thesis. During the first session, the listener was given paperwork to fill out, including a consent form (see Appendix D), a language background questionnaire (see Appendix B.1) and a race/ethnicity questionnaire (see Appendix C). All forms were approved by the Institutional Review Board (IRB) of the University of South Florida. Next, a pure-tone hearing screening was conducted, which followed the same procedures that were used for the production experiment described above. Listeners were then presented with stimuli via headphones in one example block in order to familiarize them with the task. The example block contained 12 words presented in quiet with one break halfway through.

After the example block, a level setting block was presented in order to give the listener an opportunity to request a softer or louder listening level for the remainder of the experiment. The level setting block consisted of a total of 24 words presented in noise, with a break after every six words. The listener was able to ask the investigator to adjust the listening level on each break; if adjustments were not needed, the listener was instructed to continue to the next block. None of the listeners required any level adjustments for this task; they all reported the listening level of 70 dB SPL to be comfortable.

Following the level setting block, a set of practice trials was presented to get the listener accustomed to listening in noise. The practice trials consisted of four blocks of 72 words, for a total of 288 practice trials. The first practice block was presented in quiet. The remaining blocks were presented with increasing levels of noise. For

example, the second block was presented with a SNR of 0 dB, the third block was presented with a SNR of -4 dB, and the fourth block was presented with a SNR of -8 dB. Short breaks were permitted after each practice block of 72 words. All stimuli used for the example, level-setting, and practice blocks were different from the stimuli used for the main task (stimuli from dropped talkers were used for the example, level-setting, and practice blocks).

Following the practice blocks, all listeners were presented with the test stimuli in a randomized order in individual test blocks. The test blocks consisted of six blocks of 98 words, for a total of 588 trials. Test stimuli consisted of two tokens of each of the six /bVd/ words in each speech style (clear or conversational) recorded by all 49 talkers from the production experiment, for a total of 1176 stimuli. Half of these stimuli, 588, were presented on the first day of the experiment, and the other half were presented on the second day of the experiment. All test blocks were presented in noise (-8 dB for MO/EB, -4 dB for LB). Short breaks were encouraged after each block of 98 words, and a longer break (about ten minutes) was given halfway through the experiment (after block #3). These breaks were provided to reduce fatigue. All test stimuli were randomized, and all listeners were presented with stimuli in a different randomized order. In subsequent sessions, each listener was given a set of practice trials at the start of the session to refresh their memory of the task. The second day of the experiment consisted of only the practice and test blocks. Test stimuli presented on the second day of testing included the second token of each word that was not presented on the first day of the experiment. Listeners were provided with the same instructions as the first day. The

number of stimuli and breaks was consistent with day one, as well. The instructions provided to all listeners on this task can be found in Appendix F.

As each listener completed the perceptual task, their answers and information on the correct target word were saved into a text document. These data were imported into an Excel spreadsheet, and the number of correct and incorrect responses was computed for each condition. For example, if the target word was “bid,” but the participant selected “bead,” an incorrect response was recorded in the Excel sheet, and data on the correct response and alternative chosen were saved. Listeners’ confusions (i.e., the words that were perceived rather than the target word for incorrect responses) were not analyzed for this thesis.

The number-correct data were then used to compute percent-correct data for each talker group (summed across the talkers), target word and speaking style, for each listener. Percent-correct data were also computed for each speaking style and individual talker within a group, summed across the target vowels and tokens. Prior to statistical analysis, all percent-correct data were converted to rationalized arcsine transform units (RAUs), which are more appropriate for the parametric statistical analysis than percent-correct data (Studebaker, 1985). Percent-correct data are used for the presentation of the figures.

Chapter Three

Results

To address the first two research questions, a three-way within-subjects analysis of variance (ANOVA) was used. The independent variables were the talker group (three levels: monolingual, early bilingual and late bilingual), speaking style (two levels: conversational and clear), and target word six levels: “bead, bid, bayed, bed, bad” and “bod”). The dependent variable was the RAU-transformed percent-correct data for each listener. To obtain the percent-correct scores for each listener, at each level of the independent variables, the number of correct responses was summed across the talkers within a group, with separate scores for each target vowel and speaking style. Percent correct was then computed based on the number of possible correct responses (number of possible correct responses = number of talkers in a group X 2 tokens of each target word).

All of the main effects and interactions in the three-way ANOVA were significant. The F values, degrees of freedom and p values for all of the main effects and interactions are shown in Table 3, below. Only the talker group by speaking style interaction and the three-way interaction will be discussed in detail because these are the two effects that address the first two research questions.

Table 3

Results of Three-Way ANOVA of the Effects of Talker Group, Speaking Style and Target Word

Effect	F (df)	<i>p</i> value
Main effects		
Talker group	26.75 (2,38)	<.001
Speaking style	50.06 (1,19)	<.001
Target word	21.33 (5,95)	<.001
Two-way interactions		
Talker group by speaking style	4.68 (2,38)	.015
Talker group by target word	3.67 (10,190)	<.001
Speaking style by target word	7.66 (5,95)	<.001
Three-way interaction		
Talker group by speaking style by target word	2.38 (10,190)	.011

Note. Data on F values, degrees of freedom (df) and levels of significance (p values) for all main effects and interactions in the three-way ANOVA of the effects of talker group, speaking style and target word on intelligibility of target words presented in noise to monolingual native English-speaking listeners.

The significant talker group by speaking style interaction addresses the first research goal: to compare the three talker groups in terms of their overall clear speech

benefit in noise. Percent-correct data showing the performance of each talker group on each speaking style are shown in Figure 2.

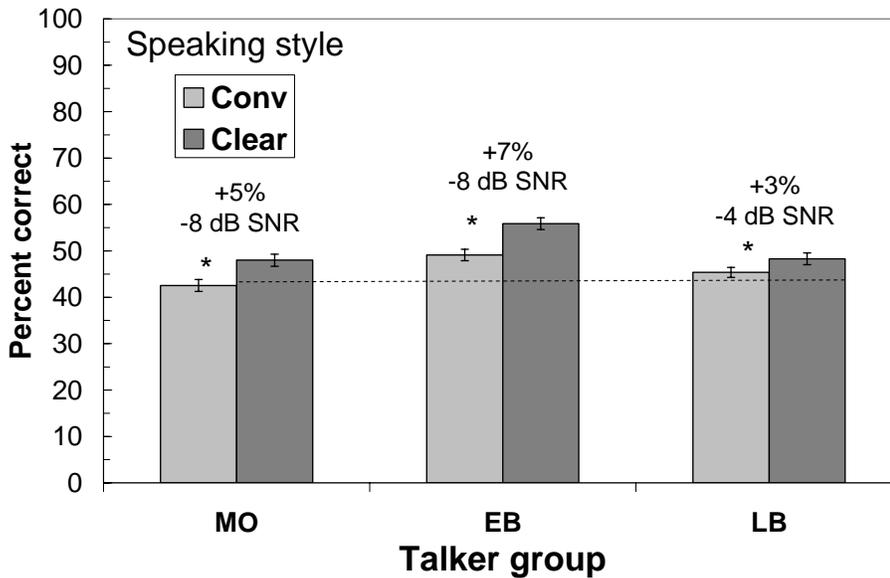


Figure 2. Mean percent-correct word-identification scores for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB), for both speaking styles. Error bars indicate one standard error of the mean. The dashed line indicates conversational style performance for the monolingual talker group. Asterisks indicate significant clear speech effects within each talker group.

As shown in Figure 2, all three talker groups—monolingual (MO), early bilingual (EB) and late bilingual (LB)—were able to produce some clear speech benefit. When comparing conversational speech to clear speech, the MO group was 5% more intelligible in the clear speech condition; the EB group was 7% more intelligible in the clear speech condition; and the LB group was 3% more intelligible in the clear speech condition. Overall, the early bilinguals performed better in both speech conditions, clear and

conversational, than the other two groups. The MO and EB groups both had a similar degree of improvement in the clear speech condition, while the LB group had a smaller degree of improvement from conversational to clear. It is important to keep in mind that the LB group was performing at a more favorable SNR of -4 dB, while the other two groups were performing at a more challenging SNR of -8 dB. As a result of this finding, one could say that the effect of being a late bilingual is equal to about a 4 dB worse SNR for vowels when compared to the other two groups. Thus, the late bilinguals were less intelligible than the monolinguals and early bilinguals.

Simple main effects post-hoc comparisons were used to explore the significant talker by speaking group interaction. First, the two speaking styles were compared for each talker group. Performance in the clear speech condition was significantly higher than performance in the conversational speech condition for each talker group ($p < .001$ for the MO and EB groups and $p = .018$ for the LB group). Thus, each group showed a significant clear speech benefit, despite the differences across groups in the size of that benefit.

Next, the performance of the MO and EB talker groups (the two groups for which the same SNR was used) were compared within each speaking style. Performance for the EB group was significantly higher than performance for the MO group in both speaking styles. This difference was highly significant for both speaking styles ($p < .001$), but was slightly larger for the clear speech style (an 8% difference between the EB and MO groups in the clear style vs. a 7% difference between the EB and MO groups in conversational style).

The significant three-way interaction addresses the second research goal: to compare the differences in the effects of clear speech across individual target vowels for each talker group. Percent-correct data showing the performance of each listener group for each speaking style are shown in Figures 3-5 below. Two of the target words are shown in each figure, with each word shown as a separate panel. Simple main effects post-hoc comparisons were used to explore the significant talker group by speaking style by target word interaction and will be discussed with the general patterns of results.

All three talker groups produced a similar size clear speech benefit for the target words “bead” (Figure 3A) and “bod” (Figure 5B). The simple main effects post-hoc comparisons showed that the clear speech benefit was not significant for “bead” for any of the talker groups (benefit = +3 to +5%); however, for “bod,” the benefit was significant for all three groups (+11 to +14%; MO: $p=.001$; EB/LB: $p<.001$). For the target word “bid” (Figure 3B), the monolinguals and early bilinguals each produced a 4% clear speech benefit (not significant); whereas, the late bilinguals performed worse in the clear speech condition, with a significant decrement of 8% ($p=.006$).

For the target word “bayed,” (Figure 4A) only the early bilinguals showed a significant clear speech benefit (+10%; $p<.001$); the other two groups showed a smaller, non-significant benefit. For the target word “bed” (Figure 4B), only the monolingual talkers were able to produce a significant clear speech benefit (+9%; $p=.01$); the early and late bilingual talkers showed no change from conversational to clear. Finally, for the target word “bad” (Figure 5A), both early and late bilingual talkers, but not the monolingual talkers, were able to produce a significant clear speech benefit (+9%; $p=.005$ for the EB talkers; +7%; $p=.041$ for the LB talkers).

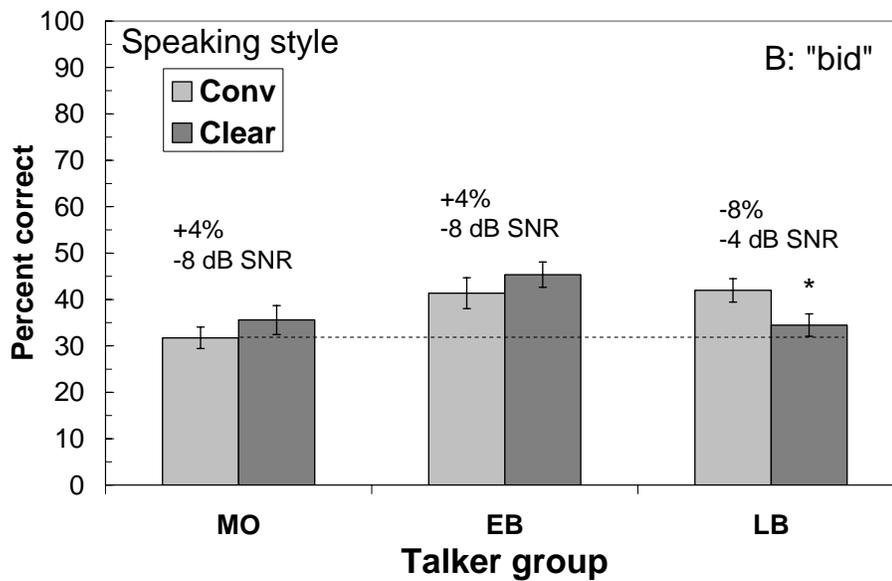
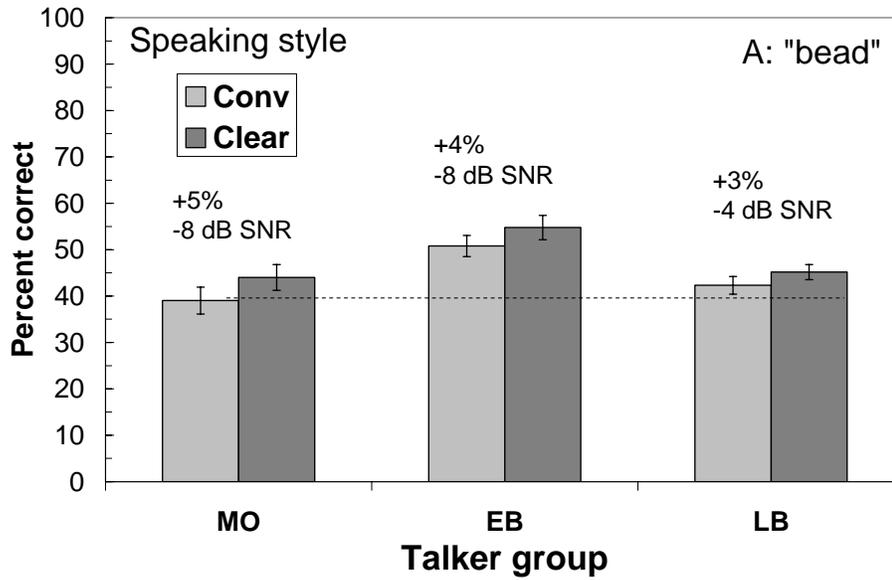


Figure 3. Mean percent-correct word-identification scores for “bead” (panel A) and “bid” (panel B) for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB). Error bars indicate one standard error of the mean. The dashed line indicates conversational style performance for the monolingual talker group. Asterisks indicate significant clear speech effects within each talker group.

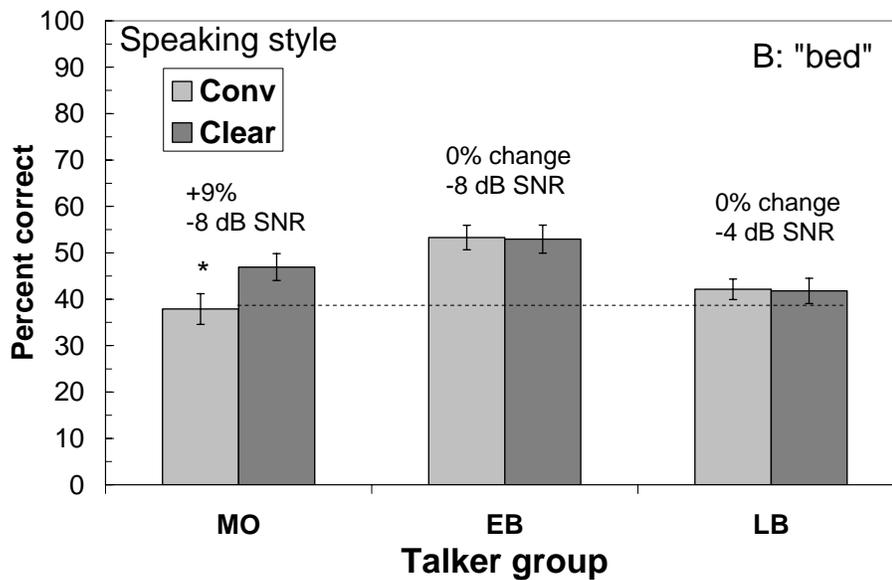
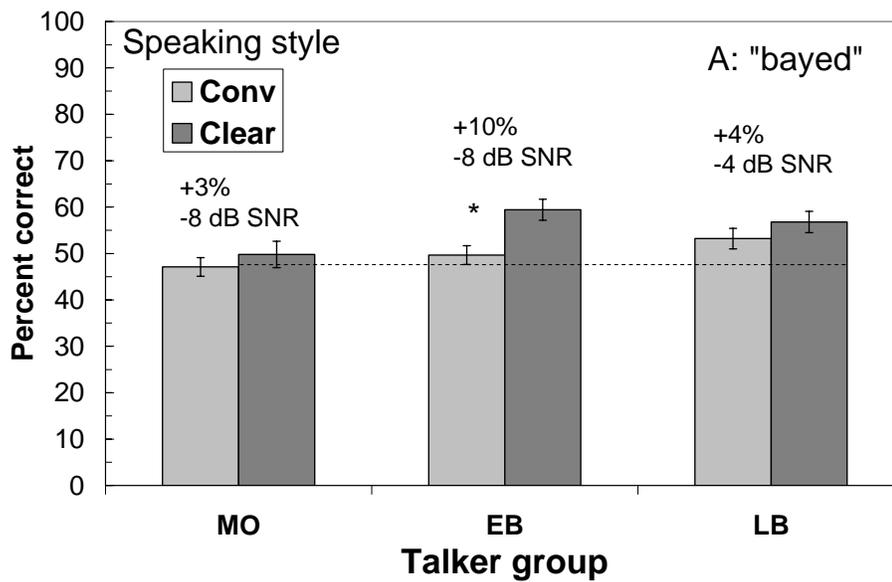


Figure 4. Mean percent-correct word-identification scores for “bayed” (panel A) and “bed” (panel B) for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB). Error bars indicate one standard error of the mean. The dashed line indicates conversational style performance for the monolingual talker group. Asterisks indicate significant clear speech effects within each talker group.

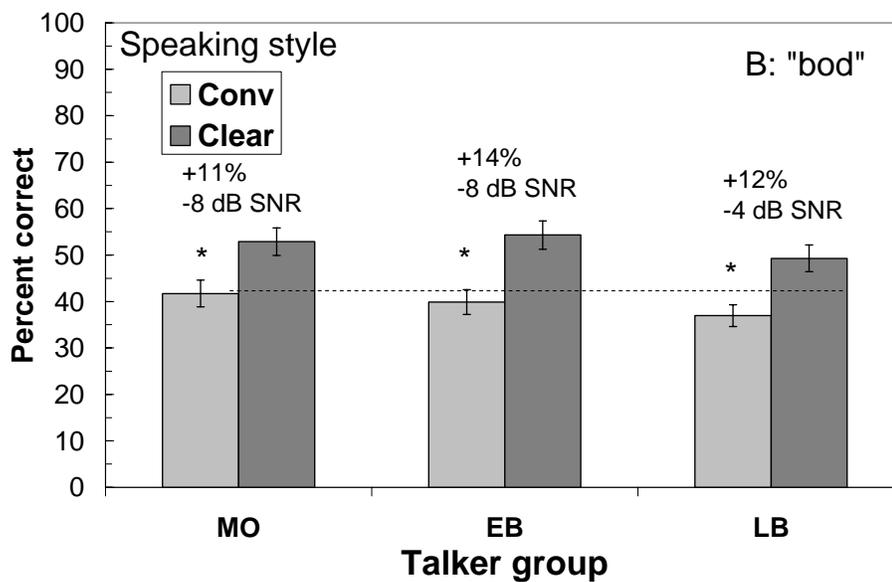
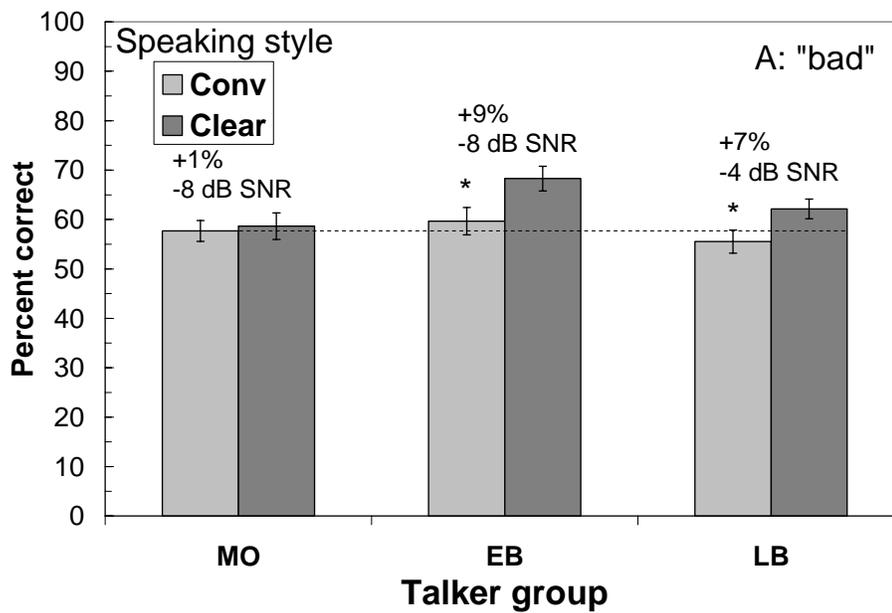


Figure 5. Mean percent-correct word-identification scores for “bad” (panel A) and “bod” (panel B) for the three talker groups: monolingual (MO), early bilingual (EB) and late bilingual (LB). Error bars indicate one standard error of the mean. The dashed line indicates conversational style performance for the monolingual talker group. Asterisks indicate significant clear speech effects within each talker group.

Next, simple main effects post-hoc comparisons were used to compare performance of the MO and EB groups (the two groups for which stimuli were presented at the same SNR) for each target word, within each speaking style. In the conversational speech style, performance of the EB group was significantly higher than that of the MO group for the target words “bead, bid” and “bed.” The two groups did not differ significantly in performance for the target words “bayed, bad” and “bod.”

In the clear speech style, performance of the EB group was significantly higher than that of the MO group for the target words “bead, bid, bayed,” and “bad.” Performance of the EB group was also higher for the target word “bed,” but the difference only approached significance ($p=.06$). The performance of the two groups did not differ significantly for the target word “bod.”

To address the third research question, the performance of the individual talkers was compared within each talker group. Percent-correct scores were computed separately for each talker within a group for each of the two speaking styles. To obtain this value, the number of correct responses for each talker was summed across the six target vowels and two tokens per vowel for each speaking style (number of possible correct responses: six vowels X two tokens per vowel = 12), with a separate score for each listener. Percent-correct data showing the performance of each talker within each listener group on each speaking style are shown in Figures 6-8. Results for each talker group are presented as a separate figure.

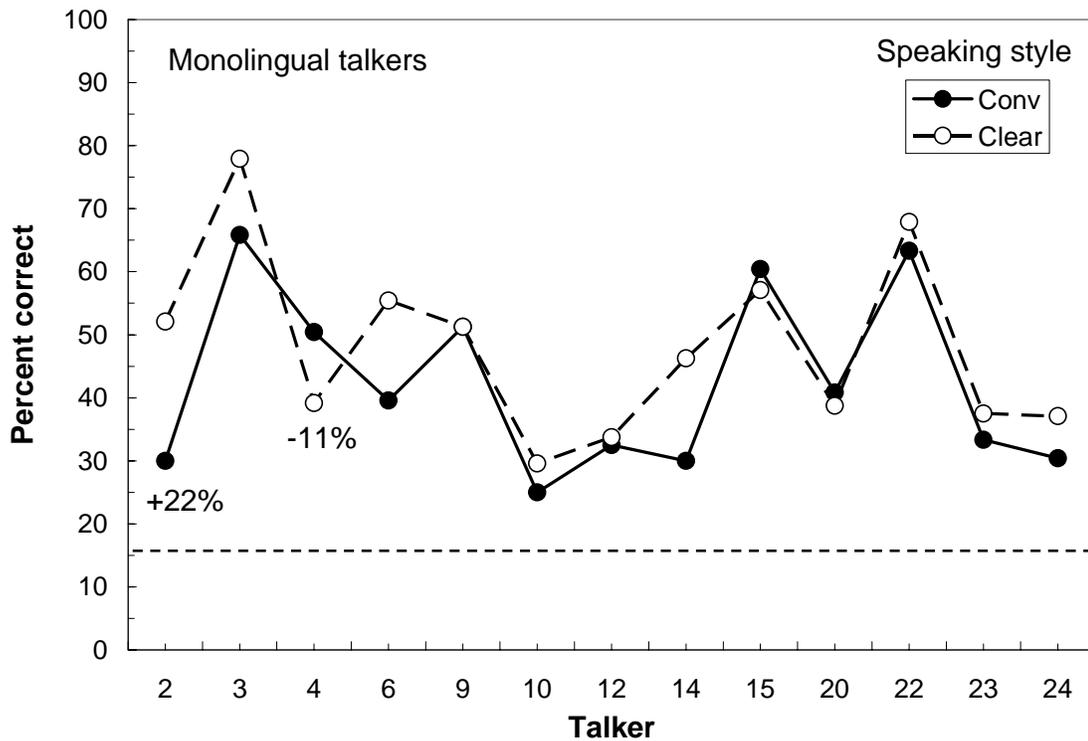


Figure 6. Mean percent-correct word-identification scores of each talker in the monolingual (MO) talker group. Performance for the conversational style (Conv) is shown by solid lines and performance for the clear style (Clear) is shown by dashed lines. The dashed horizontal line indicates chance performance (approximately 17% correct) for the six-alternative forced choice task. The difference in performance between clear and conversational speaking styles is indicated for the talkers with the greatest increases and decreases in performance from the conversational to the clear speaking style.

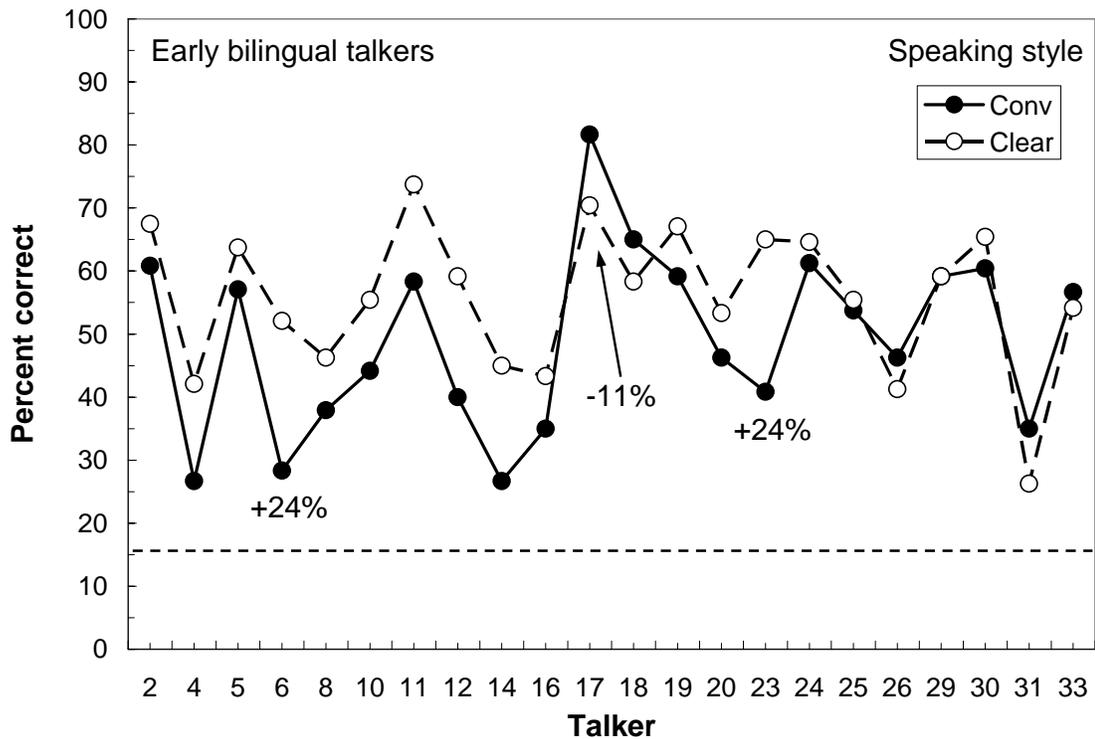


Figure 7. Mean percent-correct word-identification scores of each talker in the early bilingual (EB) talker group. Performance for the conversational style (Conv) is shown by solid lines and performance for the clear style (Clear) is shown by dashed lines. The dashed horizontal line indicates chance performance (approximately 17% correct) for the six-alternative forced choice task. The difference in performance between clear and conversational speaking styles is indicated for the talkers with the greatest increases and decreases in performance from the conversational to the clear speaking style.

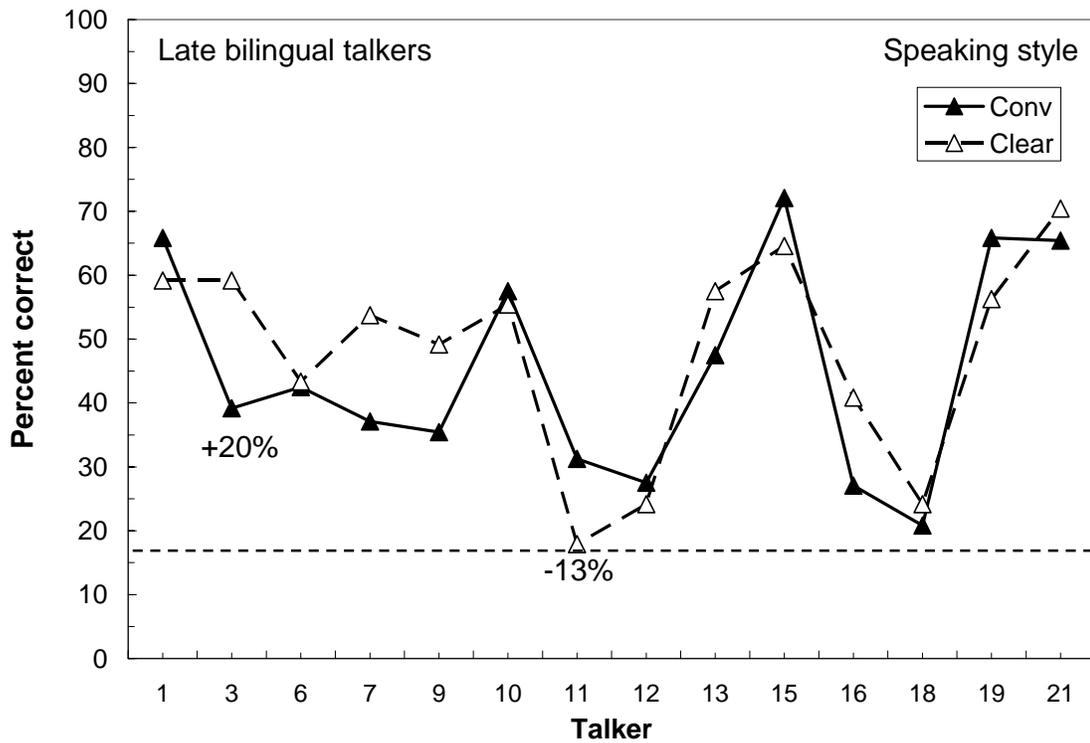


Figure 8. Mean percent-correct word-identification scores of each talker in the late bilingual (LB) talker group. Performance for the conversational style (Conv) is shown by solid lines and performance for the clear style (Clear) is shown by dashed lines. The dashed horizontal line indicates chance performance (approximately 17% correct) for the six-alternative forced choice task. The difference in performance between clear and conversational speaking styles is indicated for the talkers with the greatest increases and decreases in performance from the conversational to the clear speaking style.

As shown in Figure 6 (MO talker group), 9 of the 13 monolingual talkers (69%) showed improvement from conversational to clear speech conditions; 3 talkers (23%) performed worse in the clear speech condition; and 1 talker (8%) performed the same across speech conditions. Five talkers (38%) produced a clear speech intelligibility

benefit of more than 5%. The greatest improvement was seen for subject NA02, who improved by 22% from conversational to clear. The greatest decrement was seen in subject NA04, who was 11% less intelligible in the clear speech condition.

Con conversationally, the monolinguals ranged in intelligibility from 25% to 66% (41% intelligibility range in conversational speech). In the clear speech condition, they ranged from 30% to 78% (48% intelligibility range in clear speech).

As shown in Figure 7 (EB talker group), 16 of the 22 early bilingual talkers (73%) showed improvement from conversational to clear speech conditions; 5 talkers (23%) performed worse in the clear speech condition; and 1 talker (5%) performed the same across speech conditions. Thirteen talkers (59%) produced a clear speech intelligibility benefit greater than 5%. The greatest improvement was seen for subjects HI06 and HI23, who each improved by 24% from conversational to clear. The greatest decrement was seen in subject HI17, who was 11% less intelligible in the clear speech condition. Conversationally, the early bilinguals ranged in intelligibility from 27% to 82% (55% intelligibility range in conversational speech). In the clear speech condition, they ranged from 26% to 74% (47% intelligibility range in clear speech).

As shown in Figure 8 (LB talker group), 8 of the 14 late bilingual talkers (57%) showed improvement from conversational to clear speech conditions, and 6 talkers (43%) performed worse in the clear speech condition. Five talkers (36%) produced a clear speech intelligibility benefit of more than 5%. The greatest improvement was seen for subject LO03, who improved by 20% from conversational to clear. The greatest decrement was seen in subject LO11, who was 13% less intelligible in the clear speech condition. Conversationally, the late bilinguals ranged in intelligibility from 21% to 72%

(51% intelligibility range in conversational speech). In the clear speech condition, they ranged from 18% to 70% (52% intelligibility range in clear speech).

While there is variability amongst individual talkers in all three talker groups, it can be concluded that the monolinguals and early bilinguals performed most similarly to each other in both speech conditions; however, the early bilinguals were found to be more intelligible overall in both speech conditions. Additionally, the late bilinguals were the least intelligible in both speech conditions, with two individual talkers performing only slightly above chance.

Chapter Four

Discussion

The goal of oral communication is for the talker's ideas to be understood by the listener. Intelligibility is one way of measuring whether the message was transmitted successfully. Sometimes the talker's intent of the message is misunderstood by the listener. This misunderstanding may be the result of a decrease in the intelligibility of the message. Two barriers that can affect intelligibility are bilingualism and level of background noise. When two people are having a conversation and there is a high level of background noise, such as in a crowded restaurant, chances are that parts of the conversation will be less intelligible, and thus, a breakdown in communication may occur. Imagine this same situation, but imagine that one of the communication partners is bilingual and they are communicating in the bilingual's second language. The environmental factor of noise from the restaurant compounded by the talker factor of bilingualism may further contribute to reduced intelligibility of the message.

Some bilinguals speak their second language with a certain degree of foreign accent. This may pose a problem for these individuals in their daily lives, whether at school, at work, at home, or in the social realm. They may find that they have to repeat themselves frequently because their listeners may have difficulty understanding them. In turn, this may also lead to increased levels of frustration for both the non-native talker and the listener. For these reasons, ongoing research is needed to address the issues that present difficulty to both native and non-native talkers of the language. The present

study adds to our understanding of the problem because the smaller clear speech benefit found for late bilinguals suggests that they may lack knowledge or control of which cues to enhance when asked to speak more clearly, reflecting differences in linguistic knowledge. Understanding the degree to which bilinguals can improve intelligibility when asked to speak clearly can be useful in accent modification therapy for those bilinguals who seek it, which can help combat the problem of reduced intelligibility.

The present study addressed the issues of bilingualism and noise through the following research goals. First, a comparison across three talker groups, monolinguals, early bilinguals, and late bilinguals, was made to determine the differences in native and non-native talkers' ability to produce a clear speech benefit in noise. Second, a comparison was made of the differences in the effects of clear speech across six individual target vowels for each talker group. Third, the performance of individual talkers was compared within each group.

Summary of Results

In the present study, there were three independent variables: talker group, vowel, and speech condition. The talker group variable consisted of three levels: native English monolinguals, early Spanish/English bilinguals, and late Spanish/English bilinguals. The vowel variable had six levels: /i, I, e, ε, æ, a/. The speech condition variable had two levels: conversational and clear speech. All three talker groups were recorded while reading lists of carrier phrases. All carrier phrases were embedded with a /bVd/ word which contained the six vowel variations. Recordings were made in two different speech conditions, conversational and clear. All /bVd/ words were isolated, mixed with noise, and presented to a group of native English-speaking monolingual listeners via a six-

alternative forced-choice perception task. The dependent variable in this study was the percent of words correctly identified by the listeners.

To address the first research goal of talker group by speaking style, it was found that all three talker groups were able to produce a significant clear speech benefit, although the benefit was only about half as much for the late bilinguals as for the other two groups. It should be kept in mind also that the late bilinguals were performing at a less challenging SNR than the other two groups, due to their overall lower intelligibility level. The degree of clear speech benefit obtained was similar for the monolingual and early bilingual talker groups; however, overall, the early bilinguals were significantly more intelligible than the monolingual talkers in both the conversational and clear speaking styles.

The fact that the early bilinguals behaved similarly to the monolinguals was not a surprise. Previous research has proven that the earlier a second language is acquired, the more native-like the talker will become in the L2. Since the early bilinguals in the present study all reported learning English intensively prior to age 12 years, many of them seemed to have achieved native-like proficiency in English. The late bilinguals, on the other hand, all reported learning English intensively after the age of 15 years, which would explain why they were less intelligible overall. It was, however, an interesting finding that the early bilinguals performed significantly better than the monolinguals. This can possibly be attributed to the fact that bilinguals have had to manipulate two or more languages for most of their lives. As a result, bilinguals may have greater metalinguistic awareness, which in turn, may have enabled them to more easily manipulate the acoustic cues for the vowel studied here. Alternatively, the early

bilinguals may have had a larger vowels space, due to the need to maintain phonetic categories for both L1 and L2 vowels (cf. Flege, 1995).

To address the second research goal of talker group by speaking style by individual vowel, four general patterns were found. First, for “bid,” the monolinguals and early bilinguals better in the clear than in the conversational speech condition, although the benefit was not significant for either group. The late bilingual talkers, however, actually performed significantly worse for “bid” in the clear speech condition than in the conversational speech condition.

The second vowel pattern was shown for “bed.” Results indicated that only the monolinguals were able to produce a significant clear speech benefit. The two bilingual groups performed similar to each other in that there was no change in performance across speech conditions for both groups of bilinguals. Overall, the early bilinguals performed better than the monolinguals here, as well, however. Third, for “bad” and “bayed,” the early bilinguals showed a the largest clear speech benefit. For “bayed,” the benefit was significant only for the early bilinguals, but for “bad,” both the early and late bilingual talkers (but not the monolingual talkers) showed a significant clear speech benefit.

Finally, for “bead” and “bod,” all three groups showed a similar degree of clear speech benefit, but the benefit was only significant for “bod.”

The fact that the late bilinguals had a significant decrement for “bid” was an interesting finding. Since confusion analyses have not been completed yet, one can only speculate as to why this occurred. Most likely, the target word “bid” was confused most often with “bead.” According to Flege’s (1997) Speech Learning Model, /I/ would be considered a “new” vowel for Spanish learners of English, while /i/ would be considered

an “old” vowel. The vowel /i/ is longer in duration than /I/. On the production task, when instructed to use clear speech, the late bilingual talkers may have manipulated the duration cue by slowing down their rate of speech, and thus, lengthening the vowel. If they did this when producing the vowel /I/, they may have actually produced it acoustically more /i/-like, which would result in confusions with /i/, and thus, would explain the significant decrement in the clear speech condition for /I/. This hypothesis can be investigated by confusion analyses of the perception data and by acoustic analyses of the production data.

The finding that neither bilingual group was able to produce a clear speech benefit for /bed/ was interesting. The hypothesis that /ε/ is a “new” vowel to Spanish learners of English could possibly explain this result. The bilinguals may have had difficulty accessing the appropriate cues to produce this vowel due to a limited previous experience with it. The last vowel pattern was expected. All three talker groups were able to produce a significant clear speech benefit for “bod.” Since this vowel occurs in both English and Spanish, it was not a surprise that all three groups were able to utilize the appropriate cues to make themselves more intelligible for this vowel.

To address the third research goal of examining individual talker differences within each group across speech conditions, it was found that there was variability amongst individual talkers within each group. For the monolinguals, many of the talkers showed improvement in the clear speech conditions, but there were a few who performed worse. Among the early bilinguals, there were also many talkers who improved from conversational to clear, while others showed the opposite pattern. Finally, for the late bilinguals, a similar degree of variability in performance across talkers was found,

although overall, their clear speech benefit was smaller than for the other two groups.

There were two individual talkers who were performing slightly above chance in the late bilingual group.

The results from the third research goal were not unexpected. Every individual is unique and there may be a combination of factors that contributed to the variability found amongst talkers across all three groups. Some of these factors can include vocal quality, fundamental frequency, vocal intensity, vocal resonance, linguistic experience, etc. It is obvious that there are a number of factors that may contribute to individual variability.

Hypotheses

At the start of the present study, the investigators formulated three hypotheses. First, they hypothesized that there would be no change in vowel production across conversational and clear speech conditions for the late bilinguals for some target vowels. This hypothesis was supported to some degree. Although no talker group showed a significant clear speech benefit for all six target words, only the late bilinguals showed a significant decrement for one target word. Moreover, the average degree of benefit for the late bilinguals was only about half that of the other two groups. The late bilinguals did produce a significant clear speech benefit for two vowels /æ/ and /a/, however.

The second hypothesis was that the early bilinguals would produce a clear speech benefit in noise. This hypothesis was also supported. The early bilinguals did in fact produce a significant clear speech benefit in noise overall that was similar in size to that for the monolingual talkers. Moreover, although the clear speech benefit was only significant for target vowels /e, æ/ and /a/, performance was higher for the early bilingual talkers than for the monolingual talkers for half of the target vowels.

Lastly, it was hypothesized that the native talkers would also produce a clear speech benefit in noise. This hypothesis was supported as well. The native talkers were able to produce a clear speech benefit in noise overall. However, like the early bilinguals, this benefit was only significant for some target vowels (/ɛ, æ/ and /a/).

Comparison with Previous Research

Previous research comparing intelligibility of vowels in conversational and clear speech with monolingual talkers was been conducted by Ferguson and Kewley-Port (2002) and Ferguson (2004). Ferguson and Kewley-Port (2002) was the first study to systematically investigate vowel identification in noise in both clear and conversational speech. In Ferguson (2004), recordings were made of multiple native talkers, reading sentences containing /bVd/ words in clear and conversational speech. The words were later isolated from the sentence and embedded in 12-talker babble for presentation to native listeners in a vowel identification task. The present study is very similar to Ferguson (2004) in that /bVd/ words were produced both clearly and conversationally, their stimuli were embedded in multi-talker babble, and they presented their task to native listeners. While Ferguson (2004) addressed the linguistic factor of speaking style, the talker factor of bilingualism was not addressed. When comparing the two studies, the major differences are that the Ferguson (2004) study used only native talkers and ten target vowels, while the present study used both native and non-native talkers and only six target vowels. Ferguson found that their native talkers were able to produce a clear speech benefit of 8%. In the present study, a clear speech benefit was found for all talker groups, but it was smaller than the benefit found in Ferguson (2004). Native talkers in the present study produced a clear speech benefit of about 5%, early bilinguals produced

a 7% benefit, and late bilinguals produced a 3% benefit. Some possible reasons as to why the native talkers in the present study did not produce the same benefit as in the Ferguson (2004) study could be that there was a smaller number of talkers. The present study only had 13 native talkers, while the previous study had 41 native talkers. Individual talker differences may have had a greater impact on the present study than in the Ferguson (2004) study. Also, all of the talkers in the present study were female, while Ferguson (2004) included both male and female talkers. The factors of sample size, number of vowels studied and gender may explain why differences were found across these studies.

Bilingual research on clear speech perception has been conducted by Bradlow and Bent (2002). In their study, native English talkers were recorded reading sentences in both clear and conversational speech. The sentences were later embedded in white noise and presented to native and non-native listeners in a sentence recognition task. Results indicated that non-native listeners received a much smaller clear speech benefit in noise than did native listeners. This study paralleled the present study in that both clear and conversational speech was presented to listeners and the bilingualism factor was addressed; however, Bradlow and Bent (2002) used native talkers and native and non-native listeners, while the present study used native and non-native talkers and only native listeners. Also, Bradlow and Bent (2002) did not group their bilingual subjects by age of acquisition of English, while the present study did. Since early learners tend to be much more native-like than late learners (Rogers, et al. 2006), the question of whether a clear speech benefit differs amongst bilingual listeners from various L2 proficiency levels was not addressed by Bradlow and Bent (2002). Bradlow and Bent (2002) found that

non-native listeners received less of a clear speech benefit in noise than did native listeners. In the present study, it was found that early bilinguals produced a clear speech benefit that was similar to that produced by monolinguals, while late bilinguals produced a much smaller clear speech benefit than monolinguals. Together, the results of Bradlow & Bent (2002) and the present study suggest that late bilinguals are faced with greater challenges than early bilinguals when communicating in noisy environments due to differences in their ability both to receive a clear speech benefit when listening and to produce a clear speech benefit when speaking.

The present study can also be assessed with regard to its implications for Flege's (1995) Speech Learning Model. In Flege's model there is a distinction between "old" and "new" vowel categories for second language learners. In the present study, a comparison can be made of "old" versus "new" vowels for the bilingual talkers. For the three "old" vowels, /i, e/ and /a/, both bilingual groups were able to produce a clear speech benefit for all three of these vowels. However, for the early bilinguals, this benefit was only significant for the vowels /e/ and /a/. For the late bilinguals, this benefit was only significant for the vowel /a/. For the three "new" vowels, /ɪ, ε/ and /æ/, neither bilingual group was able to produce any change across speech conditions for the vowel /ε/. For the vowel /ɪ/, the early bilinguals produced a small, but not significant clear speech benefit and the late bilinguals actually performed significantly worse in the clear speech condition. Finally, for the vowel /æ/, both bilingual groups were able to produce a significant clear speech benefit.

Implications

Theoretically, early bilinguals seem to have the productive linguistic knowledge needed to improve intelligibility when asked to speak clearly, but the late bilinguals appear not to share this same knowledge or seem less able to access it to manipulate the relevant acoustic cues. One hypothesis is that early bilinguals may have better awareness of language due to the early need to manage two languages. It is also possible that the early bilinguals are more careful in this type of task than the late bilinguals.

Practically, the early bilinguals do not seem to be at a disadvantage when talking in noisy listening environments, in relation to the monolinguals. In fact, as mentioned previously, they actually performed better than the monolinguals. However, the late bilinguals appear to be at an increased disadvantage when talking under adverse listening conditions. Under normal listening conditions, the late bilinguals are at a disadvantage to begin with, and when noise is added to the background signal, they appear to be at an even greater disadvantage by experiencing a reduced ability to make themselves more intelligible to compensate for the challenging environment.

Clinically, the results of this research can be useful to bilinguals who wish to improve their intelligibility under adverse listening conditions. For example, a late bilingual professor, who teaches in his non-native language, may be difficult to understand by many of his students. The professor may wish to improve his intelligibility for the benefit of himself and his students. As a result, he may wish to enroll in an accent modification program at the local university, where he would be provided with therapy from a speech-language pathologist (SLP), specializing in foreign accent.

Initially, the SLP may review communication strategies with the professor. These strategies may include: slow down rate of speech, repeat and rephrase utterances, summarize at the end of each lecture to ensure the students understood the material presented, use visual aids such as a PowerPoint presentation or handouts, make sure there is nothing obstructing the mouth while talking, engage in face-to-face communication, maintain good eye contact with the students, stand in good lighting where there is not a lot of glare, and if possible, lecture in a room that has good acoustics such as carpeting and heavy curtains. It is also important to reduce the amount of background noise in the classroom in order to improve the signal-to-noise-ratio. If necessary, the professor can also use a microphone. Many of these strategies could be applied in other settings for the professor, as well.

Following communication strategies training, the SLP could then implement an accent modification program, that should be backed up by evidence-based practice such as the evidence presented in the present study, in particular information as to which of the vowels studied would be most likely to be misperceived. SLPs can provide services to their clients by training them on specific phoneme cues that are needed to improve intelligibility. Training such as this can have an enormous impact on many individuals who use their voice professionally including professors, company CEOs, customer service representatives, political candidates, actors/actresses, and regular everyday people who wish to improve their accent in order to improve their quality of life.

Limitations

Some limitations to the present study include the following: all talkers were female so it is not known how male talkers would perform on this task; therefore, this

research may not be clinically applicable to male bilinguals who wish to improve their intelligibility. Another limitation was that only six vowels of American English were studied, when there are 14 vowels in the language. It is not known which cues the talkers will manipulate for the other eight vowels. The present study only investigated clear speech effects for monolinguals and Spanish-English bilinguals. There are many bilinguals who speak languages other than Spanish, and those languages were not explored; therefore, the results of this research can only be applied directly to Spanish-English bilinguals.

In the present study, the speech produced by all talkers in the conversational style was not completely natural. Although the participants were instructed to read sentences in a normal, conversational manner, it was not true conversation since they were reading lists of carrier phrases. Normal conversation is usually full of various intonation patterns, as well as, different types of sentences (declaratives, interrogatives, exclamations, etc.). One final limitation was that the talkers were not talking under adverse listening conditions in the clear speech style. Instead, adverse listening conditions were simulated by instructing the talkers to pretend they were speaking to someone with a hearing impairment or to someone who was having difficulty understanding them. Therefore, it is not known, if the clear speech produced by the talkers is reflective of the clear speech they would actually use in a noise or other demanding listening environments.

Future Research

Future research in the near future should begin with the confusion matrices for the perception data collected in the present study. This will provide information on what types of errors the talkers are making. Next, acoustic analyses should be completed on

all of the production data in both clear and conversational styles. This will provide information on which acoustic cues, specifically, are being manipulated by the monolingual and bilingual talkers when instructed to speak clearly. This information would also be helpful in designing an evidence-based accent modification program.

Other future research could include replicating the present study to compare whether the same results are found. Replications can also include other variables such as young and old listeners, other vowels, consonants, and sentences. Other foreign languages, besides Spanish, can be investigated, as well, in order to examine whether differences exist in vowel production across languages and what those differences are.

One other possible suggestion that would be interesting would be to conduct an experimental treatment design, where late bilinguals are instructed to read sentences using clear and conversational speech at baseline; their production data can be embedded in noise and presented to a group of native listeners via a vowel identification task; next, the late bilinguals can be provided with communication strategies training and accent modification therapy; then, the late bilinguals can participate in the same experiment at baseline; and finally, their data after treatment can be embedded with noise and re-administered to the same group of native listeners in a vowel identification task, as done previously. An experiment such as this could be beneficial because it would either support or fail to support the effectiveness of communication strategies training and accent modification therapy for late bilinguals.

References

- Bilger, R. C., Neutzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research, 27*, 32-48.
- Boersma, P. & Weenik, P. (2006). Praat: Doing phonetics by computer (Version 4.4.24) [computer program]. Retrieved June 19, 2006 from <http://www.praat.org>.
- Bohn, O. S., & Flege, J. E. (1997). Perception and production of a new vowel category by adult second language learners. In A. James & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 53-73). Berlin: Mouton de Gruyter.
- Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America, 112*, 272-284.
- CoolEdit 2000 (version 1.1) [computer software] (2000). Syntrillium, Inc.: Phoenix, AZ.
- Dalbor, J. B. (1969). *Spanish pronunciation: Theory and practice* (3rd ed.). Fort Worth: Holt, Rinehart and Winston.
- Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *Journal of the Acoustical Society of America, 116*, 2365-2373.
- Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 112*, 259-271.
- Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore: York Press.

- Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437-470.
- Grosjean, F. (1989). Neurolinguistics, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3-15.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099-3111.
- Hillenbrand, J., & Nearey, T. M. (1999). Identification of re-synthesized /hVd/ utterances: Effects of formant contour. *Journal of the Acoustical Society of America*, 105, 3509-3523.
- Kent, R. D., Dembowski, J., & Lass, N. J. (1996). The acoustic characteristics of American English. In N. J. Lass (Ed.), *Principles of experimental phonetics* (pp. 185-225). St. Louis: Mosby-Year Book, Inc.
- Kewley-Port, D., Akahane-Yamada, R., & Aikawa, K. (1996). Intelligibility and acoustic correlates of Japanese accented English vowels. *ICSLP 96 (International Conference on Spoken Language Processing) Proceedings*.
- Ladefoged, P. (1982). *A course in phonetics* (2nd ed.). San Diego: Harcourt, Brace Jovanovich.
- Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41, 329-335.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of

- hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, 28, 96-103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, 29, 434-446.
- Rogers, C. L., & Dalby, J. (2005). Forced-choice analysis of segmental production by Chinese-accented English speakers. *Journal of Speech, Language, and Hearing Research*, 48, 1-17.
- Rogers, C. L., Dalby, J., & Nishi, K. (2004). Effects of noise and proficiency on intelligibility of Chinese-accented English. *Language and Speech*, 47, 139-154.
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27, 465-485.
- Small, L. H. (1999). *Fundamentals of phonetics: A practical guide for students*. Boston: Allyn & Bacon.
- Smith, D. D. (2007). *Introduction to special education: Making a difference* (6th ed.). Boston: Pearson/Allyn & Bacon.
- Studebaker, G. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 494-509.
- TDT System III [Computer hardware and software] (2001). Tucker-Davis Technologies, Inc.: Gainesville, FL.
- Tye-Murray, N. (2004). *Foundations of aural rehabilitation: Children, adults, and their family members* (2nd ed.). Australia: Thomas Delmar Learning.

United States Census Bureau (2000). 2000 U.S. Census. Retrieved July 16, 2007, from
<http://www.factfinder.census.gov>

Appendices

Social Sciences/Behavioral Adult Informed Consent

University of South Florida

Information for People Who Take Part in Research Studies

For monolingual recording and listening purposes—compensated

The following information is being presented to help you decide whether or not you want to be a part of a minimal risk research study. Please read carefully. If you do not understand anything, ask the Person in Charge of the Study.

Title of Study: *Speech perception and production by native and non-native speakers*

Principal Investigator: Catherine L. Rogers, Ph.D.

Study Location(s): PCD 3008A

You are being asked to participate because you are a monolingual native speaker of English, with normal hearing and no history of speech or hearing disorders.

General Information about the Research Study

The purpose of this research study is to examine performance of native and non-native speakers of English in both listening and speaking tasks. Audio recordings of your speech will be made and presented to listeners in two additional studies for comparison with foreign-accented speech.

Plan of Study

To participate in this experiment you must first pass a basic audiometric hearing screening. During the hearing screening, a series of short tones (beeps) of different frequencies will be played. Your task is to indicate when you hear a sound, by raising your hand. If you do not pass the hearing screening, it does not necessarily mean that you have a hearing loss, however you may wish to pursue further testing at a doctor's office or clinic. The main experiment will take place in two sessions, on two separate days. On the first day, audio recordings of your speech will be made as you repeat a set of sentences in English that will be presented to you on a computer screen. Some of the sentences will also be presented over headphones. For the audio recordings, you will speak into a microphone, connected to a digital recorder. The audio recordings of your voice will be stored on a computer. Later, portions of these recordings will be presented to listeners in two additional studies. During the main portion of the study you will be asked to produce about 95 sentences. Each will be presented on a computer screen. You will have several practice items before beginning the main task. You will be instructed in the pronunciation of any nonsense words. Following this session, you may be asked to repeat some sentences more clearly. Next, you will be asked to listen to English words. On each of the 108 trials, you will hear a single word. Your task will be to choose the word you heard from several choices. The first day's testing and completion of this form will take approximately two hours. On the second day, you will again be asked to listen to English words and to choose the word you heard from several choices. Some of the words may have been altered from their original form. Completion of the 240 trials on the second day should take about one and a half hours. On both days you will be given a short break after each 20 minutes of listening or speaking. Total time for participation will be about three and a half hours.

Payment for Participation

As a subject, you may receive either extra credit points (if you are enrolled in a participating course in a participating department) or monetary compensation for your participation. If you are to receive monetary compensation, you will be paid on an hourly basis (\$10.00/hour) for your participation in the study. If you are to receive extra credit points, you will receive 1 point for every 30 minutes of participation. You will be compensated at the conclusion of the research study sessions. If you decide to withdraw or the experimenter decides to terminate your participation, you will be compensated only for the hours of participation in the research study session prior to withdrawal/termination.

Benefits of Being a Part of this Research Study

No practical benefits of this study are likely to apply to you directly. On the other hand, by taking part in this research study, you may increase our overall knowledge of human listeners' ability to understand speech under different listening conditions. This knowledge may aid in developing improved methods of teaching English to speakers of English as a second language.

Risks of Being a Part of this Research Study

There are no known risks related to participation in this research study. The speech you will listen to will not be loud enough to be harmful under normal conditions. The level used will be about that of speech at a normal conversational level. Also, you are encouraged to notify the investigator immediately if any of the sounds are uncomfortable. If the situation is uncomfortable in any way, the experimental procedure will be changed.

Confidentiality of Your Records

Your privacy and research records will be kept confidential to the extent of the law. Authorized research personnel, employees of the Department of Health and Human Services and the USF Institutional Review Board may inspect the records from this research project.

The results of this study may be published. However, the data obtained from you will be published using a code number or will be combined with data from other people in the publication. The published results will not include your name or any other information that would in any way personally identify you. As stated above, the sentences you produce will be presented to listeners in two additional studies. Your name will not be given to the listeners. No other personal information will be given to the listeners.

Individual data will be stored and identified using a number code. All data will be stored in the Speech Perception and Production Laboratory (PCD 3008A) or in the principal investigator's office (PCD 4013). Both locations will be locked and secured. Authorized research investigators, agents of the Department of Health and Human Services and the USF Institutional Review Board may inspect your records from this research project.

Volunteering to Be Part of this Research Study

Your decision to participate in this research study is completely voluntary. You are free to participate in this research study or to withdraw at any time. If you choose not to participate, or if you withdraw, there will be no penalty or loss of benefits that you are entitled to receive. If participating for extra credit points, you will be credited for the time you participated and will receive no grade penalty for withdrawing. If participating for money, you will be compensated for the time (hours or part hours) that you participated at the rate stated above (\$10.00 per hour).

Social Sciences/Behavioral Adult Informed Consent

University of South Florida

Information for People Who Take Part in Research Studies For bilingual recording and listening purposes—compensated

The following information is being presented to help you decide whether or not you want to be a part of a minimal risk research study. Please read carefully. If you do not understand anything, ask the Person in Charge of the Study.

Title of Study: *Speech perception and production by native and non-native speakers*

Principal Investigator: Catherine L. Rogers, Ph.D.

Study Location(s): PCD 3008A

You are being asked to participate because you are a bilingual speaker of Spanish and English or speaker of English as a second language, with normal hearing and no history of speech or hearing disorders.

General Information about the Research Study

The purpose of this research study is to examine performance of native and non-native speakers of English in both listening and speaking tasks. Audio recordings of your speech will be made and presented to listeners in two additional studies.

Plan of Study

To participate in this experiment you must first pass a basic audiometric hearing screening. During the hearing screening, a series of short tones (beeps) of different frequencies will be played. Your task is to indicate when you hear a sound, by raising your hand. If you do not pass the hearing screening, it does not necessarily mean that you have a hearing loss, however you may wish to pursue further testing at a doctor's office or clinic. The main experiment will take place in two sessions, on two separate days. On the first day, audio recordings of your speech will be made as you repeat a set of sentences in English that will be presented to you on a computer screen. Some of the sentences will also be presented over headphones. For the audio recordings, you will speak into a microphone, connected to a digital recorder. The audio recordings of your voice will be stored on a computer. Later, portions of these recordings will be presented to listeners in two additional studies. During the main portion of the study you will be asked to produce about 95 sentences. Each will be presented on a computer screen. You will have several practice items before beginning the main task. You will be instructed in the pronunciation of any unfamiliar items. Following this session, you may be asked to repeat some sentences more clearly. Next, you will be asked to listen to English words. On each of the 108 trials, you will hear a single word. Your task will be to choose the word you heard from several choices. The first day's testing and completion of this form will take approximately two hours. On the second day, you will again be asked to listen to English words and to choose the word you heard from several choices. Some of the words may have been altered from their original form. Completion of the 240 trials on the second day should take about one and a half hours. On both

days you will be given a short break after each 20 minutes of listening or speaking. Total time for participation will be about three and a half hours.

Payment for Participation

As a subject, you may receive either extra credit points (if you are enrolled in a participating course in a participating department) or monetary compensation for your participation. If you are to receive monetary compensation, you will be paid on an hourly basis (\$10.00/hour) for your participation in the study. If you are to receive extra credit points, you will receive 1 point for every 30 minutes of participation. You will be compensated at the conclusion of the research study sessions. If you decide to withdraw or the experimenter decides to terminate your participation, you will be compensated only for the hours of participation in the research study session prior to withdrawal/termination.

Benefits of Being a Part of this Research Study

No practical benefits of this study are likely to apply to you directly. On the other hand, by taking part in this research study, you may increase our overall knowledge of human listeners' ability to understand speech under different listening conditions. This knowledge may aid in developing improved methods of teaching English to speakers of English as a second language.

Risks of Being a Part of this Research Study

There are no known risks related to participation in this research study. The speech you will listen to will not be loud enough to be harmful under normal conditions. The level used will be about that of speech at a normal conversational level. Also, you are encouraged to notify the investigator immediately if any of the sounds are uncomfortable. If the situation is uncomfortable in any way, the experimental procedure will be changed.

Confidentiality of Your Records

Your privacy and research records will be kept confidential to the extent of the law. Authorized research personnel, employees of the Department of Health and Human Services and the USF Institutional Review Board may inspect the records from this research project.

The results of this study may be published. However, the data obtained from you will be published using a code number or will be combined with data from other people in the publication. The published results will not include your name or any other information that would in any way personally identify you. As stated above, the sentences you produce will be presented to listeners in two additional studies. Your name will not be given to the listeners. No other personal information will be given to the listeners.

Individual data will be stored and identified using a number code. All data will be stored in the Speech Perception and Production Laboratory (PCD 3008A) or in the principal investigator's office (PCD 4013). Both locations will be locked and secured. Authorized research investigators, agents of the Department of Health and Human Services and the USF Institutional Review Board may inspect your records from this research project.

Volunteering to Be Part of this Research Study

Your decision to participate in this research study is completely voluntary. You are free to participate in this research study or to withdraw at any time. If you choose not to participate, or if you withdraw, there will be no penalty or loss of benefits that you are entitled to receive. If participating for extra credit points, you will be credited for the time you participated and will

receive no grade penalty for withdrawing. If participating for money, you will be compensated for the time (hours or part hours) that you participated at the rate stated above (\$10.00 per hour).

Questions and Contacts

If you have any questions about this research study, contact Catherine Rogers, Ph.D. at (813) 974-7423.

If you have questions about your rights as a person who is taking part in a research study, you may contact a member of the Division of Research Compliance of the University of South Florida at 813-974-5638.

Your Consent—By signing this form I agree that:

I have fully read or have had read and explained to me this informed consent form describing a research project.

I have had the opportunity to question one of the persons in charge of this research and have received satisfactory answers.

I understand that I am being asked to participate in research. I understand the risks and benefits, and I freely give my consent to participate in the research project outlined in this form, under the conditions indicated in it.

I have been given a signed copy of this informed consent form, which is mine to keep.

Signature of Participant

Printed Name of Participant

Date

Investigator Statement

I have carefully explained to the subject the nature of the above protocol. I hereby certify that to the best of my knowledge the subject signing this consent form understands the nature, demands, risks and benefits involved in participating in this study.

Signature of Investigator
Or Authorized research investigators
designated by the Principal Investigator

Printed Name of Investigator

Date

Institutional Approval of Study and Informed Consent

This research project/study and informed consent form were reviewed and approved by the University of South Florida Institutional Review Board for the protection of human subjects. This approval is valid until the date provided below. The board may be contacted at (813) 974-5638.

Approval Consent Form Expiration Date:

Revision Date: _____

Appendix B.1: Language Background Questionnaire for Monolinguals

Participant Background Questionnaire (Form A)

Name: _____ Age: _____ Address (town & state): _____

Phone (optional). Home: _____ Office: _____

Email address (optional): _____

1. Is English your first (native) language? Circle one: Yes No

1a. If you answered "No" to (1) above, list your first language here.

2. Did you speak any languages other than English while growing up (other than classroom instruction)? Circle one: Yes No

2a. If you answered "Yes" to (2) above, list those languages here _____

3. List any languages you speak other than English and rate your degree of proficiency on a scale from "1" to "5" for each (1=beginner, can't have a conversation; 5=like a native speaker):

4. Have you ever been diagnosed with a speech or hearing disorder or had speech or hearing difficulties? Circle one: Yes No

a. If you answered "yes" to (4), above, please explain in the space provided below (or on back if you need more room):

5. How long have you lived in Florida (or current state)?

6. What state were you born in and how long did you live there?

(don't answer #'s 7 or 8 if you've lived all your life in 1 state)

7. What state have you lived the longest in? _____

a. How many years did you live there? _____

8. List any other states that you've lived in for over a year (if more than 3, list top three): _____

9. On a scale from "1" to "7," rate your experience with listening to speakers with a foreign accent (1=little or no experience; 7=every day or very frequent):

Appendix B.2: Language Background Questionnaire for Bilinguals

Participant Background Questionnaire (Form B)

Name: _____ Age: _____ Address (town & state): _____

Phone (optional). Home: _____ Office: _____

Email address (optional): _____

1. How many years have you lived in your current area (town & state)?

2. Have you ever been diagnosed with a speech or hearing disorder or had speech or hearing difficulties? Circle one: Yes No

a. If you answered "yes" to (2), above, please explain in the space provided below (or on back if you need more room):

3. What language(s) did your parents speak with you? _____

a. If you answered with more than one language in (1), above, which language(s) did each parent speak with you?

4. Where were you born (give city, state, country) _____

a. How many years did you live there? _____

b. List other cities or regions you've lived in for more than one year and note number of years you lived there for each.

c. What city and country are your parents from?

Mother: _____ Father: _____

5. How old were you when you began learning English? _____

a. Why did you begin learning English? _____

6. If you moved to the United States from another country, how much did you speak English before moving here (describe years of study, if you learned English in a classroom & percent of time speaking English)?

7. If you moved to the United States from another country, how long have you lived

here? _____ years, _____ months.

8. On a typical day, what percent of your time do you spend speaking English at work? _____ % At home? _____ % Other (shopping, etc.)? _____ %

9. On a typical day, what percent of your time do you spend speaking a language other than English at work? _____ % At home? _____ %

Other (shopping, etc.)? _____ % (if more than one, answer below for each language)

10. What percent of your day do you spend with people with people who speak both (or more) languages that you do? _____ %

11. What language are you most comfortable speaking? _____
a. How much more comfortable are you in speaking that language on a scale of 1 to 5? (1=equal or nearly equal comfort; 5=much more comfortable)

12. What language are you most comfortable listening in? _____
a. How much more comfortable are you in listening in that language on a scale of 1 to 5? (1=equal or nearly equal comfort; 5=much more comfortable) _____

13. What language are you most comfortable reading in? _____
a. How much more comfortable are you reading in that language on a scale of 1 to 5? (1=equal or nearly equal comfort; 5=much more comfortable)

14. What language are you most comfortable writing in? _____
a. How much more comfortable are you writing in that language on a scale of 1 to 5? (1=equal or nearly equal comfort; 5=much more comfortable)

15. Do you think your ability in the language you are less comfortable in is still improving for any of the skills in questions 9-12? Circle one: yes no

a. If you answered yes in 13 above, indicate which abilities you believe are still improving.

Circle any that apply: speaking listening reading writing

16. What academic degrees have you earned? (list language of education for each)

17. For all languages that you speak, rate your level of ability on a scale of 1 to 5 (1=not proficient, like a child or beginner; 5=very proficient, like a well-educated native speaker) for each of the following areas:

a. Comprehension: _____

b. Fluency: (ease of expression) _____

c. Vocabulary: _____

d. Pronunciation: _____

e. Grammar: _____

Race and ethnicity questionnaire

Subject code: _____

Please put an “x” in the box for the items that apply to you.

Gender: Female

Male

Ethnicity

Hispanic or Latino. A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of the race. The term “Spanish origin” can be used in addition to “Hispanic or Latino.”

Not Hispanic or Latino.

Race (check all that apply)

American Indian or Alaska Native. A person having origins in any of the original peoples of North, Central, or South America, and who maintains tribal affiliation or community attachment.

Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent, including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

Black or African American. A person having origins in any of the black racial groups of Africa. Terms such as “Haitian” or “Negro” can be used in addition to “Black” or African American.

Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

Check here if you do not wish to provide some or all of the above information.

Social Sciences/Behavioral Adult Informed Consent

University of South Florida

Information for People Who Take Part in Research Studies For monolingual listening purposes--volunteer

The following information is being presented to help you decide whether or not you want to be a part of a minimal risk research study. Please read carefully. If you do not understand anything, ask the Person in Charge of the Study.

Title of Study: *Speech perception and production by native and non-native speakers*

Principal Investigator: Catherine L. Rogers, Ph.D.

Study Location(s): PCD 3008A; PCD3008

You are being asked to participate because you are a monolingual native speaker of English, with normal hearing and no history of speech or hearing disorders.

General Information about the Research Study

The purpose of this research study is to examine performance of native and non-native speakers of English in both listening and speaking tasks.

Plan of Study

- To participate in this experiment you must first pass a basic audiometric hearing screening. During the hearing screening, a series of short tones (beeps) of different frequencies will be played. Your task is to indicate when you hear a sound, by raising your hand. If you do not pass the hearing screening, it does not necessarily mean that you have a hearing loss, however you may wish to pursue further testing at a doctor's office or clinic. The experiment will take place in four sessions, on four separate days. On each day, you will hear up to 216 practice trials and approximately 600 trials for the main experiment task. On the first day, you will also hear 12 example trials. On each trial (example, practice, or main experiment), you will hear a word and six words will be displayed on the screen. You will choose the word that best matches the word you heard. On example trials, words will be presented in order (you will know which word is coming next). On main experiment and practice trials, words will be presented in random order (you must decide which word you heard). On both main experiment and practice trials, words may be mixed with background noise (other voices). You will be given a short break after about 5 minutes of listening and a longer break after about 15 minutes of listening. Completion of the main task and additional tasks should take about 2 hours on the first day and about 1 hour on each of the three remaining days. Total time for participation, including completion of this form should take between four and six hours.

•

Payment for Participation

You will not be paid for your participation in this study.

Benefits of Being a Part of this Research Study

- No practical benefits of this study are likely to apply to you directly. On the other hand, by

taking part in this research study, you may increase our overall knowledge of human listeners' ability to understand speech under different listening conditions. This knowledge may aid in developing improved methods of teaching English to speakers of English as a second language.

Risks of Being a Part of this Research Study

- There are no known risks related to participation in this research study. The speech you will listen to will not be loud enough to be harmful under normal conditions. The level used will be about that of speech at a normal conversational level. Also, you are encouraged to notify the investigator immediately if any of the sounds are uncomfortable. If the situation is uncomfortable in any way, the experimental procedure will be changed.

Confidentiality of Your Records

- Your privacy and research records will be kept confidential to the extent of the law. Authorized research personnel, employees of the Department of Health and Human Services and the USF Institutional Review Board may inspect the records from this research project.

The results of this study may be published. However, the data obtained from you will be published using a code number or will be combined with data from other people in the publication. The published results will not include your name or any other information that would in any way personally identify you.

Individual data will be stored and identified using a number code. All data will be stored in the Speech Perception and Production Laboratory (PCD 3008A) or in the principal investigator's office (PCD 4013). Both locations will be locked and secured. Authorized research investigators, agents of the Department of Health and Human Services and the USF Institutional Review Board may inspect your records from this research project.

Volunteering to Be Part of this Research Study

- Your decision to participate in this research study is completely voluntary. You are free to participate in this research study or to withdraw at any time. If you choose not to participate, or if you withdraw, there will be no penalty or loss of benefits that you are entitled to receive. If participating for extra credit points, you will be credited for the time you participated and will receive no grade penalty for withdrawing. If participating for money, you will be compensated for the time (hours or part hours) that you participated at the rate stated above (\$10.00 per hour).

Questions and Contacts

- If you have any questions about this research study, contact Catherine Rogers, Ph.D. at (813) 974-7423
- If you have questions about your rights as a person who is taking part in a research study, you may contact a member of the Division of Research Compliance of the University of South Florida at 813-974-5638.

Your Consent—By signing this form I agree that:

- I have fully read or have had read and explained to me this informed consent form describing a research project.
- I have had the opportunity to question one of the persons in charge of this research and have received satisfactory answers.

- I understand that I am being asked to participate in research. I understand the risks and benefits, and I freely give my consent to participate in the research project outlined in this form, under the conditions indicated in it.
- I have been given a signed copy of this informed consent form, which is mine to keep.

Signature of Participant	Printed Name of Participant	Date
--------------------------	-----------------------------	------

Investigator Statement

I have carefully explained to the subject the nature of the above protocol. I hereby certify that to the best of my knowledge the subject signing this consent form understands the nature, demands, risks and benefits involved in participating in this study.

Signature of Investigator Or Authorized research investigators designated by the Principal Investigator	Printed Name of Investigator	Date
------------------------------------------------------------------------------------------------------------------	---------------------------------	------

Institutional Approval of Study and Informed Consent

This research project/study and informed consent form were reviewed and approved by the University of South Florida Institutional Review Board for the protection of human subjects. This approval is valid until the date provided below. The board may be contacted at (813) 974-5638.

Approval Consent Form Expiration Date:

Revision Date: _____

Appendix E: Receipt for Production and Perception Experiments

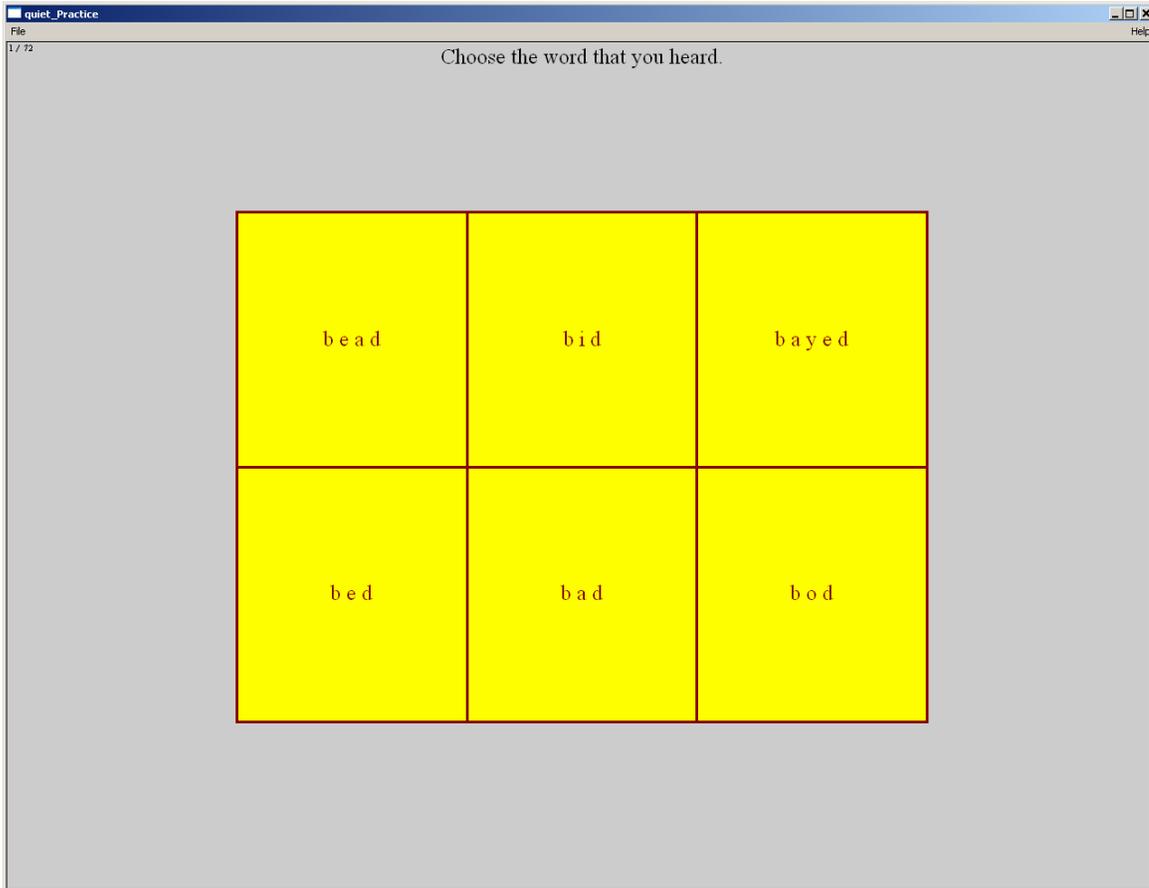
By signing this receipt, I am acknowledging that I have received the sum of \$ _____
for my participation in a University of South Florida research project, funded by the
National Institutes of Health, National Institute for Deafness and Other Communication
Disorders (NIH-NIDCD) grant #1R03DC005561-01A1, on the date noted below.

Signature of Recipient

Signature of Experimenter

Date

Appendix F: Instructions Handout for Listening Experiment



Example Instructions:

You will see a screen that looks like this. You will hear one of the following words: bead, bid, bayed, bed, bad, bod. Use the mouse to click on the word you hear. You will hear a total of 12 words with one break in between. Do you have any questions?

Level Setting Instructions:

Use the mouse to click on the word you hear. After the first 6 words, there will be background noise added to the signal. At first, it may be difficult to hear the word with noise in the background, but with practice it will get easier. Listen carefully and click the word you think you hear. If you're not sure, make your best guess. After every 6 trials, there will be more noise added, making it harder to hear the word. There will be a total of 24 trials. Let me know if the volume is too loud or too soft and I will adjust it for you. Keep in mind: when I increase the volume, the noise will also get louder and when I decrease the volume, the word will sound softer. Once we finalize a volume level that's comfortable for you, we have to keep it on that setting for the remainder of the experiment. Do you have any questions?

Practice Instructions:

Use the mouse to click on the word you hear. You will hear a total of 288 words. These words are broken up into 4 blocks. Each block contains 72 words. The first block will be presented in quiet. The second block will contain background noise. Each additional block, will contain more noise, making it harder to hear the word. If you like, take a short break in between each block. Do you have any questions?

Main Experiment Instructions (Days 1& 2):

Use the mouse to click on the word you hear. You will hear a total of 588 words presented in background noise. The noise level will remain the same for all 588 words. These words are broken up into 6 blocks. Each block contains 98 words. Take a break after each block for a few minutes. Take a longer break (~10 minutes) when you're halfway through the experiment (this will be after the third block). Do you have any questions?

Main Experiment Instructions (Days 3 & 4):

Use the mouse to click on the word you hear. You will hear a total of 588 words presented in quiet. These words are broken up into 6 blocks. Each block contains 98 words. Take a short break after each block. Take a longer break (~10 minutes) when you're halfway through the experiment (this will be after the third block). Do you have any questions?