

2009

# Spoken word recognition in quiet and in noise by native and non-native listeners: Effects of age of immersion and vocabulary size

Astrid Zerla Doty  
*University of South Florida*

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

 Part of the [American Studies Commons](#)

## Scholar Commons Citation

Doty, Astrid Zerla, "Spoken word recognition in quiet and in noise by native and non-native listeners: Effects of age of immersion and vocabulary size" (2009). *Graduate Theses and Dissertations*.  
<http://scholarcommons.usf.edu/etd/1939>

This Dissertation 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](mailto:scholarcommons@usf.edu).

Spoken word recognition in quiet and in noise by native and non-native listeners: Effects  
of age of immersion and vocabulary size

by

Astrid Zerla Doty

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Psychology  
College of Arts and Sciences  
University of South Florida

Co-Major Professor: Catherine Rogers, Ph.D.  
Co-Major Professor: Judith Bryant, Ph.D.  
Committee Member: Cynthia Cimino, Ph.D.  
Committee Member: Stefan Frisch, Ph.D.  
Committee Member: Douglas Nelson, Ph.D.

Date of Approval:  
June 30, 2009

Keywords: phonological, neighborhood, proficiency, noise, speech, perception

© Copyright 2009, Astrid Zerla Doty

## Dedication

I dedicate this dissertation to my family, especially . . .  
to Jay, for being my most adamant supporter and sharing in all the uncertainties,  
sacrifices, and joy;  
to August and Ava, whose innocence and happiness revive me daily;  
to Mom and Dad, your desire for my success and happiness made me strive for  
both;  
to Kurt, Heidi, Christine, Michele, and Eric for your encouragement and love.  
to you all, I give my deepest expression of love and appreciation.

~A dream you dream alone is only a dream. A dream you dream together is reality.  
*John Lennon (1940-1980)*

## Acknowledgements

This dissertation was completed with the help of many and I want to express my sincerest gratitude for their assistance and support. It would not have been possible without the guidance of Dr. Catherine Rogers and Dr. Judith Bryant. Dr. Rogers, my advisor and friend, for having faith and confidence in me. She managed a fine balance between giving me the freedom to pursue my ideas and reining in my imagination when it got the better of me. Dr. Bryant was always available for me and her comments were extremely perceptive and helpful. Additionally, both women are two of the nicest people I have ever known. The members of my committee, Drs. Doug Nelson, Stefan Frisch, and Cynthia Cimino generously gave their time and expertise to better my work. I thank them for their contribution and their good-natured support. I most want to thank my family. I am indebted to my parents for instilling in me the dedication and discipline to finish whatever I start and to do it well. For my mother's many home-cooked meals, for the care she gave my children during my long hours at the lab, for her tough but tender love and her tireless efforts, I am indeed thankful. My father listened to me rave about my dissertation and shared similar stories of his dissertation experience, always ending the conversations with some piece of sage advice or excruciatingly dry humor. I want to thank my siblings, Kurt, Heidi, Christine, Michele, and Eric for their support. I want to thank my husband and friend, Jay, for his love, sacrifice, and encouragement. I thank him for boosting my confidence, loving my eccentricities, and bringing joy to my life in so many ways. I also want to thank my children, August and Ava, who amaze and inspire me every day and have helped me keep perspective on what is important in life. I cannot thank them enough for showing me what it is to be a free spirit.

## Table of Contents

List of Tables	iv
List of Figures	v
Abstract	vi
Chapter 1 Introduction	1
Cross-language studies of speech perception	2
Effects of phonological characteristic on word recognition and word recall	7
The effects of cognitive demand and noise on word recognition and word recall	21
Semantic Characteristics	26
Chapter 2 Methods	31
Design	31
Participants	31
Stimuli	36
Speakers	36
Word lists	36
Recording procedures	37
Noise Mixing	39
Materials	40
Procedure	41

Chapter 3 Results	44
Word Recognition	45
Vocabulary, Language Scores, and Familiarity Ratings	50
Correlational Analyses	52
Regression Analyses	55
Digit Recall	57
Semantic Features	58
Chapter 4 Discussion	59
Phonological Neighborhood	60
Proficiency and Vocabulary Level	64
Noise	68
Digit Recall	68
Semantic Characteristics	71
Future Directions	71
References	76
Appendices	85
Appendix A: Monolingual Language Background Questionnaire	86
Appendix B: Bilingual Language Background Questionnaire	88
Appendix C: Easy and Hard Word Lists	90
Appendix D: Target Words Phonological Neighborhood and Semantic Features Data	91

Appendix E: Practice Words	94
Appendix F: Distracter Words	95
Endnote	96
About the Author	End Page

## List of Tables

Table 1	Demographic information of individual earlier-learner bilingual participants	33
Table 2	Demographic information of individual later-learner bilingual participants	34
Table 3	Demographic information for the EL and LL bilingual groups	35
Table 4	Means of Word Recognition for Easy and Hard Words by Listener Group	46
Table 5	Means of Digit Recall by Listener Group	46
Table 6	Mean percentage of times groups chose a neighbor when in error during word recognition task	49
Table 7	Vocabulary and Language Tests, Familiarity Ratings, and Demographic Variables	52
Table 8	Correlations among overall word recognition, vocabulary level (PPVT), receptive language scores (OWLS), and demographic variables (AOI and LOR) for all listeners	53
Table 9	Correlations between spoken word recognition, vocabulary, listening comprehension scores, word familiarity, and demographic variables for the EL group	54
Table 10	Correlations between spoken word recognition, vocabulary, listening comprehension scores, word familiarity, and demographic variables for the LL group	54

## List of Figures

Figure 1	Percent correct word recognition for easy and hard words for all listener groups in quiet	47
Figure 2	Percent correct word recognition for easy and hard words for all listener groups in noise	47
Figure 3	PPVT and OWLS standardized scores for all listener groups	51

Spoken Word Recognition in Quiet and Noise by Native and Non-native Listeners:  
Effects of Age of Immersion and Vocabulary Size

Astrid Zerla Doty

ABSTRACT

In spoken word recognition, high-frequency words with few neighbors and less frequently occurring minimal pair neighbors (lexically easy words) are recognized more accurately than low-frequency words with many and more frequently occurring neighbors (lexically hard words). Bradlow and Pisoni (1999) found a larger easy hard word effect for non-native than native speakers of English. The present study extends this work by specifically comparing word recognition by non-native listeners with either earlier or later ages of immersion in an English-speaking environment to that of native English speakers. Listeners heard six lists of 24 words, each composed of 12 lexically easy and 12 lexically hard words in an open-set word identification task. Word lists were presented in quiet and in moderate noise. A substantially larger easy-hard word effect was obtained only for the later learners, but a measure of oral vocabulary size was significantly correlated with performance for the non-native listener groups only. Thus, the increased easy-hard word effect for non-native listeners appears to be explained as an effect of phonetic proficiency and/or vocabulary size on the structure of the lexical neighborhoods.

## Chapter One

### Introduction

Non-native speakers of English constitute a rapidly growing minority in the United States. Many of them experience significant difficulty understanding English, especially in less than ideal listening conditions, such as the classroom or work environment. One potential source is differences in word recognition due to 1) differences in neighborhood structure; 2) greater difficulty in phonetic discrimination; and 3) language competition. Researchers have found that the effects of phonological word neighborhood on the word recognition abilities of bilinguals are greater than those found for monolinguals (Bradlow & Pisoni, 1999; Imai, Walley, & Flege, 2005). A neighborhood is a collection of words that are phonetically similar to a given target word (i.e., they sound similar) and is composed of two parts: 1) the number and degree of confusability of words in the neighborhood, referred to as ‘neighborhood density’, and 2) the frequencies of the neighbors in language use, called ‘neighborhood frequency’ (Luce & Pisoni, 1986).

The extent to which speech recognition is influenced by both phonological word neighborhood and increasing cognitive demand has yet to be investigated in either monolingual or bilingual populations. The present study compares the word recognition performance of monolinguals and bilinguals under conditions of increasing cognitive load, using words that vary in phonological neighborhood characteristics. Note that the term “recognition” refers to its traditional use in the speech perception literature, which is typically considered “perceptual identification” in memory literature.

Word recognition depends in part on the intelligibility of the stimuli, which is the degree to which something is capable of being understood. It is recognized that intelligibility of the stimuli and adults' word recognition in particular have been shown to be influenced by many variables, such as the listeners' familiarity with various aspects of the signal (e.g., the speaker, the accent of the speaker, the topic), the selection of stimuli, the task, the context, the rate of presentation of the stimuli, the listening conditions including the presence of noise and the level of variability in the stimuli, and the degree of cognitive demand required in the listening task and/or simultaneous tasks. The variables of interest in this study include phonological word neighborhood, age of immersion (AOI) of the bilingual listeners, presence of noise, and cognitive demand (in this case, a digit-recall task). In order to gain an understanding of how these specific variables may influence the recognition of speech, the discussion begins with an overview of non- native speech perception. Following this, the influences of first (L1) and second languages (L2) on each other in cross-language speech studies are considered. Next, the effects of stimulus and lexical characteristics on word recognition and recall are discussed. Finally, the effects of cognitive demand manipulations on recognition and memory for synthetic speech are explained, and parallels are drawn between findings from those studies and findings for native and non-native speech perception in noise.

#### *Cross-language studies of speech perception*

Cross-language studies of speech perception have shown that adults are language-specific perceivers. That is, although they are able to differentiate easily the phonetic categories of their native language, perception of non-native phonetic contrasts is, in general, more difficult. Phonetic contrasts are pairs of sounds in a language that differ

along a given dimension, such as voicing. This language-specific pattern of performance is not due to loss in auditory sensitivity to the acoustic features that differentiate non-native contrasts (Best, 1995). Rather, it reflects the attunement of selective perceptual processes to the acoustic-phonetic information that is linguistically relevant in the native language (Strange, 1999). However, a growing body of evidence shows that this result does not hold for all listeners, for all phonetic distinctions, or all task conditions (Flege & Hillenbrand, 1984; Rochet, 1995; Strange, 1992; Strange, Bohn, Trent, & Nishi, 2004). Clearly, an understanding of the variables used in speech perception experiments is crucial to understanding these differences and designing future studies. Also, given that selection of the languages in a cross-language investigation is based on phonetic and phonological characteristics, an understanding is necessary of the way that first and second languages influence and compete with one another. For example, if testing the discrimination of a non-native contrast for Spanish than for Italian speakers, the targeted L2 sounds may be assimilated into one phonological category for the Spanish speakers and into two phonological categories for the Italian speakers. This would presumably happen because of the closer similarity of the L2 sounds to L1 sounds for the Spanish compared to the Italian speakers. These different assimilation patterns would be predicted to result in better discrimination by the Italian speakers (Best, 1995). Moreover, researchers may test non-native speakers who are unfamiliar with the target distinction or, alternatively, who are learners with varying degrees of proficiency in the language from which the distinction is drawn and thus have varying experience with the target distinction.

According to Flege's Speech Learning Model (SLM) (1995), the process of equivalence classification prevents category formation for L2 sounds that are perceived as similar to L1 sounds. The SLM has four basic tenets relevant to this discussion: 1) the processes used in learning the L1 sound system, including category formation, do not atrophy at adolescence as asserted by the critical period hypothesis; rather, they remain functional throughout one's lifetime, 2) phonetic categories are the long-term language-specific memory representations of speech sounds, 3) the phonetic categories acquired for L1 sounds will eventually come to reflect properties of both L1 and L2 sounds that are realizations of each category, and 4) bilinguals must work to maintain contrast between those L1 and L2 phonetic categories just as monolinguals must maintain distinctness among all L1 sounds. Moreover, formation of a phonetic category implies the perceptual ability to identify a wide range of phones as being the same despite auditorily detectable differences among them along dimensions that are not phonetically relevant, as well as the ability to distinguish the multiple exemplars of a category from realizations of other categories, even in the face of non-critical commonalities. As the perceptual dissimilarity between an L2 sound and the closest L1 counterpart increases, the probability of new category formation also increases. Further, the SLM asserts that the earlier in life L2 learning commences, the smaller the perceptual distance that is needed between the L1 and L2 sound for L2 category formation to occur (Flege, 1995). Even if a new category is formed for an L2 sound, however, there is no guarantee that the category structure or weighting of acoustic cues will be the same as for monolingual speakers of the target language, according to the SLM (Flege, 1995). Thus, a mismatch may exist between a bilingual listener's perceptual expectations for a target L2 sound and the sound that is

actually produced by native speakers of the target language (Imai, Walley & Flege, 2005). This mismatch might explain reductions in L2 learners' spoken word recognition accuracy as compared to monolinguals, especially in conditions of noise or other degradation to the speech signal (Imai et al., 2005).

Another claim of the model is that when a category is not formed for an L2 sound because it is too similar to an L1 counterpart, the L1 and the L2 categories will assimilate, leading to a merged category. The sounds in this merged category should eventually come to resemble each other in production. Alternatively, when a new category for an L2 sound is established, it may dissimilate from neighboring L2 (or L1) sounds to preserve phonetic contrast of these sounds, which exist in a common phonological space. Support for these hypotheses comes from Flege, Schirru, and MacKay (2003), which examined the production of the English /eɪ/ vowel by native Italians who differed according to age of learning. Early bilinguals were found to produce English /eɪ/ with significantly more formant movement than native English speakers. The exaggerated movement of /eɪ/ produced by the early group was attributed to the dissimilation of a new phonetic category they had formed for English /eɪ/ from Italian /e/. Many of these speakers may have established a new category for English /eɪ/ (which is produced with less formant change) and produced it with more movement than is typical for English in order to make it distinct from their Italian /e/. On the other hand, the undershoot of movement observed for the late bilinguals in the study was attributed to their failure to establish a new category for English /eɪ/, which led to the merger of the phonetic properties of English /eɪ/ and Italian /e/ through the mechanism of category assimilation.

The Speech Learning Model also states that a bilingual's phonetic representation of a target speech sound may be based on different features or feature weights than those of a monolingual speaker of the L2. Support for this hypothesis comes from Flege, Bohn, and Jang (1997). Two sets of synthetic continua (11 spectral steps created by changing the first and second formants plus 3 temporal steps) were used. Subjects had to identify the vowel in one set of continua as either "beat" or "bit" and as either "bet" or "bat" in the other set of continua. Vowel duration influenced the native English subjects' identifications primarily for vowels at the middle of the continua where spectral cues were insufficient to define a vowel's identity unambiguously. When identifying members of the "bet-bat" continuum, the experienced non-native speakers (who came from various L1s) made more use of spectral cues than did the inexperienced non-native speakers. They also made less use of temporal cues. In this case, the experienced non-native speakers (also from various L1s), but not the inexperienced non-native speakers, resembled the native English speakers because they were using the spectral and temporal cues in a similar way as the native English listeners.

Despite these broad categories of "experienced" versus "inexperienced", it should be noted that the age of L2 acquisition, the degree of exposure to the language, and experience with the L2 seem to be factors that may heavily determine the relationship between perception of the first and second languages and may contribute to changes in these perceptual abilities over time. The perception of one's second language might also differ according to the class of sounds, the acoustic and perceptual correlates of these classes, and to contextual effects. Although all of these variables did not need to be controlled in the present experiment, careful consideration was given to the listeners'

proficiency level with English, and particularly age of immersion, because experience has been shown to affect perceptual abilities in the L2.

*Effects of phonological characteristics on recognition and recall*

One issue faced when investigating spoken word recognition is the structural relations among the phonological patterns of words in the mental lexicon. In addition to the contextual and stimulus factors that affect a word's intelligibility, there are lexical factors that may increase or decrease the probability or speed with which a listener will correctly identify a spoken word. In fact, it has been argued that the process of word recognition relies on accurate discrimination among competing lexical items (Luce & Pisoni, 1998). Thus, understanding the structural organization of words in memory and how these relations influence word recognition and lexical access is crucial to understanding how these factors may influence perception by L2 learners (Luce & Pisoni, 1998).

According to the Neighborhood Activation Model (Luce & Pisoni, 1998), the number of similar competitors that a word has and their relative frequency in the language can have both inhibitory and excitatory effects on lexical access. The claim is that spoken words are recognized in the context of phonologically similar words activated in memory; a spoken word activates a set or "neighborhood" of similar sounding words in memory, which then compete for recognition. A similarity neighborhood is defined as a collection of words that are phonetically similar to a given target word. A similarity neighborhood is composed of two parts: 1) the number and degree of confusability of words in the neighborhood, referred to as 'neighborhood density,' and 2) the frequencies of those neighbors, called 'neighborhood frequency.' A neighbor of a given target word

is one that differs from the target word by a one phoneme addition, substitution, or omission. For example, some neighbors for the word “sat” would be “stat, rat, sit, sap,” and “at.” The model proposes that the frequency of a given word, the size of the word’s neighborhood, and the frequency of the words within that neighborhood will determine the probability of that word being selected over its closest phonological neighbors.

The effects of phonological neighborhood on word recognition and word recall are particularly interesting. In spoken word recognition tasks, numerous studies have supported the predictions of the Neighborhood Activation Model (Goldinger, Luce, & Pisoni, 1989; Luce, 1986; Luce & Pisoni, 1998; Luce, Pisoni, & Goldinger, 1990; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985). For example, in a perceptual identification task, words with low-density neighborhoods were found to be identified in noise with greater accuracy than those from high-density neighborhoods. According to the Neighborhood Activation Model, the poorer identification of words from high-density neighborhoods is a consequence of their having more competitors so activation of the target must reach a higher level to overcome competition.

Sommers (1995) provided further support for the Neighborhood Activation Model. He found that identification accuracy of easy words was similar for young and older adults, but identification of hard words was significantly worse for older adults whose identification accuracy was 15% lower for the hard words than for the easy words, compared to younger adults, for whom the difference was only 7%. Sommers argued that the older adults may have more than just overall reduced auditory abilities; they may also have less ability to discriminate the sound patterns in the speech signals, especially from among phonetically similar neighbors. The easy-hard word effect was also

disproportionately greater for older listeners when the task demands were increased by switching from single to multiple talkers, which may suggest an influence of greater processing demand in addition to an effect of auditory abilities in the differences obtained between older and younger listeners.

Unlike the findings from recognition studies, the effects of neighborhood characteristics on word recall seem to differ from study to study and task to task. Generally, studies agree that there is better recall for high-frequency words (Allen & Hulme, 2006; Goldinger et al., 1991; Roodenrys et al., 2003). However, some studies have found better recall for words with low-frequency neighborhoods (Goh & Pisoni, 2003; Goldinger et al., 1991), whereas others have found better recall for words with high-frequency neighborhoods (Roodenrys et al., 2003). Similarly, some have found better recall for words from small neighborhoods (Goh & Pisoni, 2003; Goldinger et al., 1991) and others have found the opposite (Allen & Hulme, 2006; Roodenrys et al., 2003). The differences seem to stem from the fact that some studies considered all three variables together (word frequency, neighborhood frequency, and neighborhood density) while others considered each variable separately. Also, authors of these studies used different cut-offs for determining “high” versus “low” frequency of target word and the frequency of its neighbors, as well as the density of the neighborhoods.

For example, Goh and Pisoni (2003) used word sets that differed on neighborhood density and frequency, but were equated for word frequency. They found that recall was better for words from small, low-frequency neighborhoods than words from large, high-frequency neighborhoods. The researchers argued that there is less lexical competition

among similar sounding traces for words from small neighborhoods, which leads to less confusion among the candidates for reconstruction.

Using words that differed systematically in word frequency and neighborhood size, Allen and Hulme (2006) found better recall for high-frequency words and those from large neighborhoods compared to low-frequency words and those from small neighborhoods. In fact, the words from large neighborhoods were recalled more accurately even though they were *perceived* less accurately. Thus, it appears that the recall differences between words from large and small neighborhoods do not depend upon differences in how well these words are perceived. However, the recall advantage for high-frequency over low-frequency words may depend in part on the greater ease of perceiving high-frequency words. The authors suggest that it is the semantic representations that account for the differences in recall between words from large and small neighborhoods.

Low-frequency words in high-density, high-frequency phonological neighborhoods (i.e., words that occur relatively infrequently in the language and have many similar sounding neighbors that occur relatively often in the language) are predicted to be recognized less quickly and accurately than high-frequency words from low-density, low-frequency neighborhoods. Thus, the former are termed “hard,” whereas the latter are deemed “easy” (Luce & Pisoni, 1986). In summary, ‘easy’ words are those that occur frequently in the language and have relatively few phonetically similar neighbors that are relatively low frequency. The ‘hard’ words, on the other hand, occur less frequently in the language and have many phonetically similar neighbors that are relatively high in frequency.

Goldinger et al. (1991) selected word sets that were easy to identify (high-frequency words from sparse, low-frequency neighborhoods) and hard to identify (low-frequency words from dense, high-frequency neighborhoods) and used them in a serial recall task. Results showed better recall performance for the easy-to-identify words than the hard-to-identify words. Unfortunately, it is not possible to draw any conclusions from this study about the influence of neighborhood characteristics on verbal short-term memory performance because neighborhood characteristics were confounded with word frequency. Further, the researchers assert that the degree of confusability for given words only conveys information about the listener's internal lexicon and the relative accessibility of its component words.

Other studies provide supporting evidence for a link between speech perception processes and recall of spoken words (Luce et al., 1983; Paris et al., 2000). These studies suggest that, when the encoding of words becomes difficult, memory performance for these words declines. For example, Roodenrys et al. (2002) assessed immediate memory for word sets differing in frequency, neighborhood size, and average word-neighborhood frequency. When they considered just word frequency and neighborhood size, they found recall better for high frequency words and for words from large neighborhoods. When word frequency and neighborhood frequency were manipulated, they found that recall was better for high-frequency words and words from high-frequency neighborhoods. Finally, neighborhood size and neighborhood frequency manipulations revealed better recall for words from high-frequency neighborhoods and for words from large neighborhoods compared to small. The effects were explained in terms of word frequency (the easy to perceive [high frequency] words were recalled more

accurately than the hard-to-perceive [low-frequency] words. The researchers argued that memory was better for the high-frequency words because their representations in long-term memory are more accessible or better specified than those of low-frequency words. Interestingly, words were more likely to be intruded upon by a neighbor if they were low frequency, had many neighbors, and if the average frequency of the neighbors was high (Roodenrys et al., 2002). Typically, when a neighborhood intrusion occurred the intruding neighbor was higher in frequency than the presented word, which seems to support the predictions of Neighborhood Activation Model. Roodenrys et al. (2002) argue that phonological information in LTM plays an active role in recall in STM tasks, which helps explain the recall advantage for high-frequency words compared to low-frequency words.

In order to explain why words from large neighborhoods were recalled better than words from small neighborhoods, Roodenrys et al. (2002) suggest that the finding reflected the role of speech-production processes (e.g., retrieval of the speech motor programs for words that have to be articulated) in immediate memory tasks, but not speech-perception processes. These experiments therefore appear to provide evidence counter to the idea that word recall depends on a reintegration process which involves speech-perception mechanisms. The deleterious effects of a large neighborhood on word recognition seems to happen because the listener is required to select a word from among a large number of competitors. On the other hand, the facilitative effects of a large neighborhood on recall seem to happen because the neighbors provide support by keeping the word in active rehearsal longer than a word with few neighbors. From the existing data it appears that the effects of word frequency and neighborhood size on recall

are robust, but the effect of neighborhood frequency is small and inconsistent across experiments.

Although the studies discussed above have been limited to monolingual participants, a few studies have explored the contributions of lexical characteristics on spoken word recognition for non-native listeners. Bradlow and Pisoni (1999) investigated the combined effects of talker-, listener-, and item-related factors on isolated word recognition. The researchers had ten monolingual English-speaking talkers record both “easy” and “hard” lists of words at three different rates (slow, medium, and fast). The authors selected these words so as to differ according to three lexical characteristics. The easy words occurred more frequently in the language, their mean neighborhood density (the number of phonetic neighbors) was lower than those of the hard words, and the mean neighborhood frequency (the mean frequency of the neighbors) of the easy list was lower than that for the hard list.

Further, the frequency counts from the Brown Corpus of printed text<sup>1</sup> (Kucera & Frances, 1967) were used to examine the words to determine that the easy list words had a significantly higher mean frequency of usage in the language than did the hard list words (185.24 with a range of 36-895 versus 4.21 per million with a range of 1-35, respectively). Second, the words on the easy list were selected so that their mean neighborhood density (the number of phonetic “neighbors”) was lower than that of the words on the hard list (13.34 neighbors with a range of 3-19 versus 26.96 neighbors with a range of 21-39, respectively). Bradlow and Pisoni (1999) used the definition by Greenberg and Jenkins (1967) of a neighbor of a given target word as one that differed from the target word by a one phoneme addition, substitution, or omission. Finally, the

words on the easy list were selected so that their mean neighborhood frequency (i.e., the mean frequency of usage of the neighbors of the target word was lower than that of the words on the hard list (37.50 per million with a range of 2.33-79.67 versus 282.2 per million with a range of 87.22-1066.59, respectively). Further, the familiarity of each of the words was assessed and all were judged to be highly familiar to native-English-speaking adults. That is, they all received a rating of at least 6.25 on a 7-point scale, with 1 being lowest familiarity and 7 being highly familiar (Nusbaum, Pisoni, & Davis, 1984). In summary, the easy words are those that occur frequently in the language and have few phonetically similar neighbors that are mostly low-frequency. The hard words, on the other hand, occur less frequently in the language and have many similar neighbors that are mostly high in frequency.

Each of the listeners in the study by Bradlow and Pisoni (1999) heard the full set of 150 words spoken by a single talker at a single rate: they heard lists of words and were required to type the words they thought they heard on a computer keyboard. Note that in order for a participant to recognize a hard word, he or she had to discriminate among a large set of alternatives and, necessarily, needed to be able to make finer phonetic distinctions among words at the segmental level because the hard words had more similar sounding neighbors that were also more frequent in the language, relative to the easy words. Overall, recognition scores were significantly higher for the easy words. The authors argued that this effect of lexical discriminability resulted from the listeners' knowledge of the sound-based structure of the lexicon. In general, the results for monolinguals from this experiment replicate those of previous studies (Luce, 1986; Luce & Pisoni, 1998; Luce et al., 1990; Pisoni et al., 1985) and support the assumptions of the

neighborhood activation model of spoken word recognition; word recognition takes place within the context of the mental lexicon and, therefore, is influenced by other phonetically similar words (Luce & Pisoni, 1998).

In a second experiment, Bradlow and Pisoni (1999) found that the easy/hard word effect was greater for non-native listeners than for native listeners. In this experiment, the listeners again heard a word over headphones and typed what they heard into a computer keyboard. Two separate lists were used, one produced by a single talker and the other produced by multiple talkers. Within each list, half of the words were easy and half were hard. For both the native and the non-native listeners, the overall percent correct was higher for the single-talker condition and for the easy words. Native listeners recognized words with greater accuracy than non-natives. However, the difference in percent-correct word recognition between the easy and hard words (i.e., the easy-hard word effect) was several times greater for the non-native than for native listeners. Perhaps, as is consistent with theories of non-native speech perception (Best, 1995; Flege, 1995), non-native listeners have greater difficulty recognizing words that require perception of fine phonetic detail for discrimination because they may not have acquired all the native cues or do not have the same cue weighting as native speakers. Moreover, Bradlow and Pisoni (1999) asserted that, because non-native listeners have more difficulty with hard words than easy words, just as do native listeners, their results support the idea that the non-native speakers develop lexicons of their second language by employing the same sound-based organizational principles as native listeners.

Additionally, the authors administered a measure of word familiarity in order to assess the familiarity of the non-native listeners with the target words. For this task,

participants used a 7-point scale to rate their familiarity with a list of English words presented on a computer screen. Because the hard words occur less frequently in the language, one possibility the authors considered was that the non-natives were simply less familiar with the hard words and therefore were unable to recognize them accurately. The pattern of familiarity ratings given by the non-natives paralleled those of the native listeners: higher familiarity ratings were assigned to the easy words and lower familiarity ratings were given to the hard words. Generally, compared to the native listeners, the non-native listeners rated themselves as much less familiar with the hard words, and this was reflected in their recognition scores as well. When familiarity was controlled, however, by using only words rated as highly familiar to both native and non-native listeners in the analysis, a stronger easy/hard word effect for the non-native listeners than for the native listeners was still observed. Thus, although part of the non-natives' difficulty in recognizing hard words might have stemmed from their lack of familiarity with the words, familiarity alone does not fully account for the effect, suggesting that decreased discrimination of fine phonetic detail or other factors may also play a role.

Furthermore, a test of subjective familiarity of the target words may not fully reflect non-native listeners' lexicon. First, the non-natives might have recognized those words in spoken form but have more experience and familiarity with the words in print form than in spoken form. Garlock, Walley, and Metsala (2001) describe familiarity as encompassing two constructs: experienced frequency and age-of-acquisition. The authors give an example using the word "cartoon." This word, they argue is acquired early by most children, but it may not be encountered all that frequently by either children or adults. On the other hand, the word "cartilage," they maintain, is encountered

later in life, but may be used frequently by individuals in certain professions, such as doctors. The authors argue that high-frequency words overlap with other words on a segmental basis more often than do lower-frequency words, and these neighbors tend to be high-frequency as well. This means that the neighborhood density and frequency characteristics contribute substantially to perceived word familiarity. Furthermore, a rating of an “8” given by a native English speaker may be very different than a rating of “8” given by a bilingual who learned English late in life. Also, it is not only the familiarity with the target words that is of concern. Rather, the listeners’ knowledge of the words in the neighborhood of the targets is also of interest because the number of neighbors of a target word known by the subject could dramatically alter the structure of that neighborhood. For example, listeners with smaller vocabularies may know fewer of the low-frequency neighbors of easy words but more of the high-frequency neighbors of the hard words, thereby increasing the easy-hard word effect.

Spoken words are recognized by native listeners in the context of other words in the mental lexicon, and words requiring fine phonetic discrimination (i.e., hard words) are more difficult to recognize than words that do not require a high level of phonetic discrimination (i.e., easy words) (Luce & Pisoni, 1998). This can be especially true for non-native speakers (Best, 1995; Flege, 1995). However, in word recognition tasks, listeners are doing more than just discriminating among phonemes. They must discriminate among lexical items. Thus, it seems unlikely that spoken word recognition is accomplished solely by phonetic discrimination, rather, the stimulus input may activate a number of similar acoustic-phonetic representations and recognition must necessarily involve discrimination among lexical items (Luce & Pisoni, 1998). Therefore, it is

reasonable to assume that non-native listeners would have greater difficulty discriminating among hard words than among easy words because the hard words may be less familiar to them. Further, hard words would have more high frequency neighbors that might also be more familiar to the learner. Thus, the relative structure of the neighborhood for a bilingual might differ substantially and systematically from that of a native speaker. Such differences in neighborhood structure, based on vocabulary size, may explain how the size of the easy-hard word effect might increase for non-native listeners relative to native listeners. That is, easy-word neighborhoods may be effectively smaller for some non-natives because the target words themselves are likely to be known to them and fewer of the neighbors are likely to be known than for monolinguals. Hard-word neighborhoods on the other hand may be of similar size for both native and non-native listeners, but the words themselves less may be less familiar to the non-natives. Thus, the relative difference in neighborhood size could be increased for non-natives with substantially smaller vocabularies.

No studies to date have investigated the extent to which vocabulary size may contribute to the processing of speech by bilinguals in conditions of increased cognitive demand. As suggested by Goldinger et al. (1991), the effects of neighborhood characteristics on word recognition convey information about listeners' internal lexicon and the relative accessibility of a given word and its neighbors. Therefore, it seems reasonable to gather information from both monolingual and bilingual participants that will give insight into their internal lexicons, such as measures of receptive vocabulary size and listening comprehension.

Oral receptive vocabulary size is a measure of interest for several reasons. In the literature on child language acquisition, receptive vocabulary size has been shown to be a strong predictor of performance on both phonetic discrimination and phonological (non-word repetition) tasks. Authors of these studies speculate that a larger vocabulary size requires the child to pay greater attention to fine phonetic detail, resulting in more adult-like category formation (Majerus, Poncelet, Greffe, & Van der Linden, 2006; Walley, 1993). As described by Walley (1993), some believe that children's lexical processing is more holistic than segmental at the outset. As the vocabulary grows, it is argued, children begin a segmental restructuring of their lexical representations which allows for more phonetically detailed and efficient storage. Moreover, studies have shown that children's short-term memory performance, as measured by digit-span and non-word repetition tasks, has a strong positive correlation with vocabulary development (Majerus, Poncelet, Greffe, & Van der Linden, 2006). In a related study, although with college-aged subjects, Lewellen et al. (1993) used three measures to separate participant groups in their investigation of how differences in subjects' lexical familiarity influenced their word recognition and lexical access. They gathered data on word familiarity, vocabulary level, and language experience. Lexical familiarity was assessed by having participants rate on a 7-point scale the familiarity of 450 words that were selected from Webster's Pocket Dictionary and had familiarity ratings from a previous study (Nusbaum, Pisoni, & Davis, 1984). Based on their results, the researchers argued that IQ measures did not provide insight into the underlying cognitive processes involved in lexical access; rather they found that participants who differed in rated familiarity of the target words also differed in processing efficiency. They assumed that participants with higher scores on

the familiarity ratings, vocabulary test, and language experience questionnaire had larger lexicons and, therefore, could activate more candidates for recognition than could individuals with smaller lexicons.

Based on these issues, the first goal of this study is to compare the size of the easy-hard word effect in the recognition of spoken words by three listener groups: monolingual, earlier-learning non-native, and later-learning non-native.

As yet, however, the AOI of non-native listeners have not been considered in studies that have investigated the easy/ hard word effect or cognitive demand using the pre-load technique. Spoken word recognition by non-native speakers depends on vocabulary development in the target language, yet I am unaware of any study that has investigated the easy/hard word effect for non-native speakers that has also measured the participants' target language vocabulary level. Imai et al. (2005) defined proficiency as the degree of accentedness of the non-native speakers as measured by native listeners. They later correlated such factors as number of years of English-language study with degree of accent. Likewise, Bradlow and Pisoni (1999) performed only correlational analysis of factors such as age of English study onset, number of years of English study, and number of years in an English environment, but neither study directly measured the vocabulary level of their non-native participants. Therefore, I obtained a measure of target-language vocabulary development of the non-native speakers in order to investigate the relation between vocabulary level and word recognition. Theoretically, it was not necessary to control for first language since the effects of vocabulary size should be present in the pairing of any L1 and L2, and I was not focusing on specific phonemes.

*The effects of cognitive demand and noise on word recognition and word recall*

One potential explanation for the difficulty bilinguals experience listening to their L2 speech is that speech processing demands greater attentional resources even for proficient bilinguals than for monolinguals (Rogers et al., 2006). These differences may not be seen in quiet or undemanding conditions when attentional resources are plentiful. Investigation of the effects of bilingualism on the perception of speech presented under a range of listening conditions is important because the effects of bilingualism on listeners' perception may combine with the effects of adverse environmental listening factors in ways different from those for monolinguals. Research on synthetic speech intelligibility may offer insights to the experience of bilinguals because synthesized speech, like native speech for second-language learners or non-native speech for native listeners (Imai, Walley, & Flege, 2005), may not match a listener's expectancies for all cues. Thus, similarity effects may be in play during word recognition leading to similar perceptual effects. Below, a brief discussion of studies of perception of synthetic speech is provided in order to consider the potential processing parallels for native speakers listening to synthetic speech versus non-native persons listening to speech in their second language.

Pisoni and Koen (1981) found that monolingual listeners' word recognition performance on the Modified Rhyme Test (MRT) decreased more in noise for synthetic speech than for natural speech, even though performance in quiet was similar for both synthetic and natural speech. Likewise, Koul and Allen (1993) looked at the effects of noise on the intelligibility of synthetic and natural speech and found that decreasing signal-to-noise ratios had more deleterious effects on synthetic speech, although the patterns of errors were similar for both.

Paris, Thomas, Gilson, and Kincaid (2000) found that when linguistic cues (e.g., prosody, syntax, and semantic cues) were manipulated or eliminated in sentences, immediate recall of both synthetic and natural speech declined. Using the phonetically-balanced Harvard sentences (sentences that avoid high-predictability, too frequent use of one word, and for which phoneme frequency matches that of English), participants heard four different kinds of utterances: normal (with prosodic and contextual cues), no prosody (normal sentences with no prosody), no context (semantically anomalous sentences with prosody), and unstructured (unrelated words with no prosody).

The semantically analogous sentences were created by rearranging the words in the sentence. Additionally, the sentences with semantic context were not highly predictable and, in all speech modes, any within-word or lexical prosody remained. To create the “no prosody” and “unstructured” stimuli in the natural speech condition, individually recorded words were concatenated into strings. Participants were then required to immediately repeat what they heard. Overall, intelligibility and recall were better for natural speech than synthetic speech and for natural sentences with prosodic cues than those without. Interestingly, removing the prosody from the synthetic speech did not cause a further decrement in immediate recall than did the synthesis itself. Paris et al. (2000) suggested that the prosodic cues present in synthetic speech systems are not helpful to the listener, so removing them causes no additional decrement in intelligibility. The researchers argued that when these cues are not modeled correctly extra burden is placed on working memory that can exceed its capacity. Listeners’ attention, they contend, is drawn towards more superficial acoustic information and is directed away from deeper linguistic analyses. They argued that as intelligibility decreases, context

becomes increasingly important because listeners must depend on other sources of information for accurate word recognition. Context may be used as a “compensatory mechanism” that listeners use when intelligibility is degraded or poor, as with synthetic speech, and by extension non-native speech. Thus, prosody is only helpful when the overall intelligibility is relatively good, as in synthetic speech or like non-native speakers in the case of second language learning.

In investigating these issues, Paris, Gilson, Thomas, and Silver (1995) also found that performance on text comprehension tasks was better for natural voices as contrasted with synthetic and for easy than for hard passages. They argued that stimulus encoding and comprehension processes share a common pool of resources. If listening to synthetic speech requires that a greater proportion of cognitive resources be allocated to analyzing the initial-acoustic structure of the signal, the researchers assert, fewer resources are then available for comprehending and processing the semantic content.

One explanation of these findings is that decoding the acoustic and phonetic characteristics of synthetic speech may require more cognitive effort than decoding natural speech. This may be due to the relatively small number of acoustic cues present in synthetic speech than in natural speech, which is redundant and contains many cues that may help to specify a particular phoneme. With less redundancy in the acoustic signal, the listener has fewer converging sources of evidence regarding the identity of the phoneme or word in question and thus may have a more difficult time differentiating the target word from phonologically similar neighbors. It is hypothesized that the reduction in redundancy of acoustic cues in synthesized speech leads to more effortful processing of the speech, which may go unnoticed in conditions of quiet and when task demand is

low (Logan et al., 1989). When noise masks portions of the acoustic cues or processing demand is increased by other factors, however, the effects of the greater demand placed on the system by the synthetic speech are seen.

Pisoni, Nusbaum, and Greene (1985) hypothesized that perception of synthetic speech requires more cognitive effort than perception of natural speech for both words and non-words. This means not only that lexical retrieval that is more difficult, but also that the extra processing effort appears to be related to the process of extracting the acoustic-phonetic information from the signal. They reasoned that synthetic speech requires more short-term memory capacity and should interfere with other cognitive processes because it imposes greater capacity to process it. The consequences might mean that listeners who are trying to encode an impoverished signal, such as that found in synthetic speech, speech presented in noise, or non-native speech, could perform worse on simultaneous or subsequent cognitive tasks. Likewise, Ralston, Pisoni, Lively, Greene, and Mullennix (1991) found that on-line processing, as assessed by word monitoring and sentence-by-sentence listening was worse in all tasks with synthetic speech than for natural speech. In line with the reasoning offered by Pisoni et al. (1985), these authors suggested that poorer comprehension is due in part to the greater encoding demands required for the perception of synthetic speech.

In addition to the difficulties encountered in the perception of synthetic speech relative to natural speech, there seem to be even more deleterious effects found when the listener's task requires increased capacity demands. For example, Luce, Feustel, and Pisoni (1983) compared recall for synthetic speech and natural speech using a memory pre-load paradigm: subjects were visually presented with zero, three, or six digits and

then heard a list of words in either synthetic or natural speech. Subjects were instructed to recall the digits in the exact order and then recall as many of the words as they could. Results showed that fewer listeners were able to recall the digits accurately when they were followed by synthetic speech than by natural speech; participants also showed poorer free recall and poorer ordered recall for digits and words in the synthetic speech condition. Luce et al. (1983) argued that these differences in recall performance between synthetic and natural speech occur because synthetic speech has fewer redundant acoustic cues than natural speech, leading to impoverished representations in short-term memory. These impoverished representations mean that short-term memory has to work harder to maintain the signals in memory. Additionally, rehearsal of the digits in short-term memory may be interrupted by the greater encoding effort necessary for synthetic speech than for natural speech. The researchers argued that degraded input may require spare capacity in short-term memory, thus supporting the proposal that decrements in recall for degraded stimuli are the result of both encoding difficulties and short-term memory limitations.

If the difficulty encountered in processing synthetic speech results from differences between listeners' expectations of acoustic cues and the acoustic cues encoded in the signal, then it seems reasonable to assume that the same type of difficulty may arise for bilingual listeners, who may not have acquired all of the cues used by native listeners or who may weight these cues differently from native listeners (Flege, 1995; Imai, Walley, & Flege, 2005). That is, if the perception deficits encountered by non-native listeners are due to encoding difficulties at early processing stages, then there should be measurable increases in the demands placed on the resources available in short-

term memory for the non-native listeners, relative to native listeners. Another factor that may contribute to increased processing demand for bilinguals is the need to suppress the non-active language during processing of the active language in order to decrease interference (Grosjean, 1997). Taken together, the two factors of mismatch between phonetic expectations and input and the need to suppress the non-active language may result in substantially greater processing demands for bilingual listeners. Another goal of the study is to directly investigate the effects of cognitive load (noise and memory load) on speech processing by bilinguals to help confirm or disconfirm these hypotheses and, if they are found to be true, allow for some estimate of the magnitude and conditions of the increase in processing demand for bilinguals relative to monolinguals.

#### *Semantic Characteristics*

In addition to considering the phonological neighborhood characteristics, semantic network characteristics of words may also affect the access one has to words in recognition tasks. No previous studies of the effects of phonological word characteristics on word recognition have taken into account the potential effects of these characteristics. Whereas phonological neighborhood characteristics provide information about the relationships among words based on their sound patterning, semantic network characteristics provide information about the relationships among words based on their meanings (Nelson, McEvoy, & Schreiber, 1998). Three indices were considered: cue set size, connectivity, and concreteness. The first index considered, cue set size, refers to the number of different cued associates for a particular word. It is calculated by presenting individuals with a word and then counting the number of different responses or targets given by two or more participants in a given sample (Nelson, McEvoy, & Schreiber,

1998). Cue set size provides a relative index of the set size of a target word by giving a reliable measure of how many strong associates it has. In the words used for this study, this index did not differ between easy and hard word lists. Connectivity is an index of the average associate-to-associate connectivity among the associates of the cue and of the target. In other words, it indicates the density and level of semantic association within cues and targets. This measure differed between the easy and hard words in this study, with the hard words being less connected than the easy words. These findings appear to be contrary to the Neighborhood Activation Model (Luce & Pisoni, 1998). The phonologically hard words are less semantically connected and thus, more easily accessed due to less competition relative to the easy words. Thus, an effect in which the hard words are more difficult to recognize cannot be attributable to the difference in semantic connectivity. Alternatively, perhaps having a high degree of connectivity does not have the same consequences for word recognition as does having a large number of phonologically similar neighbors. Having lots of neighbors may make recognition for a specific word relatively difficult compared to a word with few neighbors because the neighbors compete with each other based on the way they sound to a listener. On the other hand, having a high degree of semantic connectivity could work to support the activation of a particular word because the semantically associated words work to prime the target word for recognition. The final semantic index considered, concreteness, which is a measure of the ease with which a word can be imagined as measured on a scale from 1-7, did differ between easy and hard word lists. This effect also seems to go against the predictions of the Neighborhood Activation Model (Luce & Pisoni, 1998) because the phonologically hard words are more concrete, and presumably more easily

accessed than the easy words. Therefore, an effect in which the hard words are more difficult to recognize relative to the easy words could not be easily attributable to the difference in concreteness. It should be noted here that a limitation of this study is that not all the words had data points for each of the semantic features, especially the hard words. For example, only 62 of the easy words had values for connectivity, while the hard words only had 34. Therefore, it is difficult to speculate on what any findings regarding the effects of semantic characteristics on word recognition accuracy might mean, as they should be interpreted with caution. This was purely an exploratory analysis and provides direction for future study.

Based on these issues, the first goal of the present study is to compare the size of the easy hard word effect in recognition of spoken words by three groups of listeners: monolingual English speakers, high-proficiency bilinguals, and low-proficiency bilinguals.

The second goal of this study is to determine the effects of increasing cognitive demand (in this case increasing number of digits to be recalled) on speech recognition and working memory during speech perception tasks to determine whether the effect was greater for non-native than for native listeners. As noted earlier, in their study on recall of synthetic versus natural words, Luce et al. (1983) found that the pre-load memory technique placed increased demands on the encoding and/or rehearsal processes in short-term memory when the participants were simultaneously engaged in another task that also required short-term memory capacity. This decrement in performance was worse for synthetic speech than for natural speech. Further, the stimuli in the present study consisted of the easy and hard words used by Bradlow and Pisoni (1999). In their study,

easy words had higher intelligibility than hard words, and this was especially true for the non-native listeners. They argued that the ability to make the fine acoustic-phonetic distinctions required to discriminate the hard words is a skill that develops with knowledge of the sound-based system of the language.

The findings of this study should have important implications for the teaching and assessment of non-native speakers. Specifically, working to increase the vocabulary level of second-language learners may indirectly help improve their recognition and subsequent comprehension of spoken material. As their L2 vocabulary grows, bilinguals' ability to make fine phonetic distinctions also appears to get better. Typically, classrooms are noisy and distraction-filled environments, and the task of comprehending a lecture is made more difficult when the subject matter is advanced. That holds true when all individuals are native speakers, but may become exaggerated when non-native listeners are involved. Clearly, we need more studies to elucidate the factors that have the greatest impact on intelligibility, especially for non-native talkers and listeners.

The specific research objectives were (1) to compare the effect of phonological neighborhood characteristics on word recognition between native listeners and non-native listeners, (2) to examine the effect of non-native listeners' age of immersion and vocabulary level on recognition of words and recall of digits in noise and in quiet, (3) to compare the effects of increasing number of digits to recall on the recognition of easy and hard words for native and non-native listeners in quiet and noise, and (4) to explore the effects of semantic characteristics of words on their recognition.

It was predicted that the easy/hard word effect on recognition would be greater for non-natives, especially under conditions of increased digits to recall. More

specifically, since it is argued that the ability to discriminate and subsequently encode hard words is a skill that develops with knowledge of the language (i.e., vocabulary level), I predicted that the early learners (EL) listeners in this study would have better recognition of the hard words compared to the later learners (LL) listeners. Finally, the effects of vocabulary level and noise were predicted to have an effect on the recall of digits such that the LL group would experience more deleterious effects of noise on their ability to recall the digits than would the EL group. This would result from the additional short term memory capacity needed to recollect the digits, which would subsequently leave less capacity for the encoding of words in the recognition task. There are no predictions regarding the effects of the semantic characteristics of the words on word recognition since I did not have data for all the words nor did I control for the factors. Instead, the analyses were exploratory in nature with no a-priori expectations.

It was hypothesized that vocabulary level would predict word recognition, which supports the premise that increasing vocabulary level may result in greater attention to fine phonetic detail. The more words one has in his or her vocabulary, the more necessary it becomes to be able pay attention to fine phonetic detail in order to make distinctions among them. It was further hypothesized that positive correlations between the vocabulary level of the non-native listeners and their word recognition scores would suggest that lexical development precedes and influences phonological knowledge of the L2 or a bi-directional or interactionist theory in which lower-level, phonological knowledge and higher level, lexical knowledge influence one another.

## Method

### *Design*

The experimental design is a mixed model with three digit recall conditions (0, 3 or 6 digits), two types of words to be recognized (easy versus hard), and noise (quiet versus noise) as within subjects variables. Proficiency (monolingual, earlier-learner non-native, later-learner non-native) varied between subjects. The dependent variables were number of words correctly recognized and the number of digits correctly recalled.

### *Participants*

Two groups of listeners participated in this experiment. Thirty-six monolingual English speakers (MO) born in the United States comprised the first group. According to self-report, they did not have spoken or written fluency with any language besides English (see Appendix A). The listeners in the other group consisted of sixty non-native speakers whose second language is English. This group was divided into 36 earlier and 24 later age-of-immersion (AOI) categories (see Tables 1 and 2) based on the participants' age of immersion in an English-speaking environment and other information gathered via the language-background questionnaire (see Appendix B). Theoretically, it was not necessary to control for first language (L1) because the effects of vocabulary size, noise, and number of digits to be recalled should be present in the pairings of any L1 with any L2. Listeners were between the ages of 18 and 50 years. Fifty years was chosen as the upper limit in age because, beyond this point, age-related hearing loss and age-related decreases in cognitive processing abilities are more likely to occur. The

listeners included both males and females (M=15, F=81). They were recruited from the University of South Florida Departments of Psychology and Communication Sciences and Disorders and the English Language Institute. Participants were compensated with extra credit points or were paid for their participation either with cash or gift certificates. Listeners were screened to exclude those with a history of speech, language, or hearing disorders. Potential participants were also required to pass a pure tone hearing screening prior to their participation. Native listeners did not have a strong regional accent as judged by the investigator, a native English speaker.

The earlier learners (EL) were those who were immersed in an English-speaking environment at age 10 or earlier, rated themselves as relatively balanced in proficiency in their L1 and L2 in a variety of contexts, and, according to a screening by the experimenter had at most a mild foreign accent. The later learners (LL) were those who were immersed in an English-speaking environment at age 14 or later, rated themselves as dominant in their L1 in a variety of contexts, and had a moderate to strong degree of foreign accent in the experimenter's judgment. The cut-off ages for the EL and LL groups, although relatively arbitrary, were selected because they provide a good separation between the groups in terms of age of immersion in the L2. As shown in Table 3, the EL and LL groups differed significantly in the reported percent of time spent speaking English at home and the reported amount of time spent speaking their L1 with others. They also differed in how they rated their proficiency in their L1 and L2. The EL bilinguals gave themselves significantly higher ratings than the LL bilinguals in the areas of comprehension, fluency, vocabulary, and pronunciation in English. On the other hand, the LL bilinguals gave themselves significantly higher ratings than the EL bilinguals in

the areas of comprehension, fluency, vocabulary, pronunciation, and grammar in their L1.

The groups rated themselves similarly in grammar in English.

Table 1. Demographic information of individual earlier-learner bilingual participants.

Age	AOI	L1	Country of Origin	LOR	Age	AOI	L1	Country of Origin	LOR
20	2	German	Germany	0.2	21	7	Spanish	US	21
18	5	Tagalog	Philippines	13.5	19	5	Spanish	US	19
24	10	Arabic	Egypt	14.2	19	5	Gujarati	US	19
20	5	Vietnamese	US	20	18	6	Creole	US	18
22	7	Spanish	Puerto Rico	17	20	4	Spanish	Dominican Republic	16
20	3	French Creole	Canada	16.5	18	10	Spanish	Cuba	7.5
18	4	Vietnamese	US	18	19	5	Spanish	US	19
18	9	Hindi	India	9.5	18	5	Urdu	Pakistan	13
20	5	Spanish	Mexico	19	24	1	Spanish	US	24
18	2	Spanish	Cuba	15	19	7	Spanish	US	10.8
30	1	Greek	US	30	20	4	Spanish	US	19
19	8	Urdu	Pakistan	11.5	19	6	Creole	US	5
20	5	Spanish	US	20	19	6	Tagalog	Philippines	3
20	5	Spanish	US	20	20	8	Spanish	US	19
19	6	Spanish	US	19	23	4	Creole	US	23
21	2	Creole	US	21	19	4	Spanish	US	19
20	4	Spanish	US	20	22	9	Spanish	Puerto Rico	10
20	5	Serbian	Serbia	12.3	27	5	Spanish	US	7.3

Table 2. Demographic information of individual later-learner bilingual participants.

Age	AOI	L1	Country of Origin	LOR
25	21	Japanese	Japan	4.3
25	20	Japanese	Japan	5
19	18	Spanish	Colombia	1
20	19	Portuguese	Angola	1
18	17	Spanish	Colombia	0.75
23	15	Creole	Haiti	9.25
18	14	Russian	Russia	5.5
22	14	Serbian	Bosnia	10.5
20	15	Spanish	Peru	4.75
38	28	Spanish	Colombia	10
29	20	Spanish	Nicaragua	10
22	17	Creole	Haiti	4.25
21	14	Spanish	Colombia	9
34	19	Spanish	Puerto Rico	14.3
20	19	Polish	Poland	1.25
22	14	Spanish	Peru	12
23	16	Spanish	Colombia	7
22	14	Spanish	Colombia	8.6
23	16	Albanian	Albania	7.5
49	14	Bulgarian	Bulgaria	13
23	14	Spanish	Cuba	9.3
21	14	Spanish	Colombia	7
27	18	Serbo-Croatian	Serbia	9.6
22	14	Japanese	Japan	8

Table 3. Demographic information for the EL and LL bilingual groups. Means are presented with standard deviations and ranges in parenthesis. Ratings are based on a scale from 1-5 from 1 (not proficient) to 5 (very proficient).

	Earlier Learners	Later Learners
Chronological Age	20.28 (2.59; 18-30)	24.42 (7.04; 18-49)
Age of Immersion **	5.25 (2.29; 0-10)	16.83 (3.36; 14-28)
Length of Residence **	15.81 (6.21; 0-24)	7.2 (3.84; 1-14.3)
% of time spent speaking English at home *	55.36 (30.88; 0-100)	36.63 (38.42; 0-100)
% of time spent speaking English at work	82.60 (33.44; 0-100)	72.95 (36.37; 0-100)
% of time spent speaking English in all other situations	70.15 (31.68; 0-50)	65 (30.32; 10-100)
% of time spent with speakers of their L1 *	40.19 (27.43; 0-90)	59.5 (32.29; 0-100)
Comprehension in English **	4.58 (.65; 3-5)	3.88 (.68; 3-5)
Comprehension in L1 **	4.35 (.61; 3-5)	4.83 (.48; 3-5)
Fluency in English **	4.61 (.64; 3-5)	3.71 (.64; 3-5)
Fluency in L1 **	4 (.93; 2-5)	4.88 (.45; 3-5)
Vocabulary in English **	4.28 (.74; 2-5)	3.63 (.65; 3-5)
Vocabulary in L1 **	3.75 (.87; 2-5)	4.58 (.72; 3-5)
Pronunciation in English **	4.44 (.81; 2-5)	3.25 (.85; 1-4)
Pronunciation in L1 **	4.06 (.95; 2-5)	4.92 (.28; 4-5)
Grammar in English	4.25 (.87; 2-5)	3.96 (.75; 3-5)
Grammar in L1 **	3.47 (1.18; 1-5)	4.54 (.78; 3-5)

\*= significant difference between groups at  $p < .05$

\*\*= significant difference between groups at  $p < .005$

### *Stimuli*

*Speakers.* Two monolingual speakers of English recorded lists of words. The speakers were 24- and 26-year-old women who were recruited from the University of South Florida and were judged to have no strong regional dialect by the experimenter.

*Word lists.* In the main production task, the speakers read aloud 144 words from a list provided by the investigator. All the words came from the stimuli used by Bradlow and Pisoni (1999) as described previously with a few modifications. Bradlow and Pisoni's (1999) easy and hard word sets consist of 75 easy and 75 hard words each, but, the present study used only 72 words (see Appendix C and D) from each list for two reasons. First, a multiple of 12 words was needed to fit the design of the study. Second, the word lists used by Bradlow and Pisoni (1999) overlapped on all three lexical characteristics. By omitting three words from the easy list ("fool", "wash", and "was") and three from the hard list ("main", "wrong" and "white"), there was a more defined separation between the lists. For example, "fool" and "wash" have quite low target frequencies. Easy words should have lower neighborhood frequency, but "wash" is about one standard deviation above the mean for the easy word list. Easy words should also have lower neighborhood density, yet "fool" was about one standard deviation above the mean for the easy words on this measure. "Was" is an outlier for frequency, even though in the expected direction, and also has an extremely low neighborhood density. For the hard words, "main", "wrong" and "white" are all relatively high in target frequency and are the only words to overlap in target frequency with the frequency of the "easy" words. Additionally, "wrong" is actually on both lists which was an admitted mistake on Bradlow and Pisoni's (1999) part. It also better fits the criteria for the easy words. For "white" the neighborhood density is actually lower than the mean for the easy words

(hard words should have higher density); for "main" the neighborhood frequency is substantially below average for the hard words. In summary, the easy words are those that occur frequently in the language and have few phonetically similar neighbors that are mostly low-frequency. The hard words, on the other hand, occur less frequently in the language and have many similar neighbors that are mostly high in frequency.

In addition to considering the phonological neighborhood characteristics, semantic network characteristics of the easy and hard word lists were examined. It should be noted that data on these variables was not available for many of the words. The first index, cue set size, refers to the number of different cued associates for a particular word (Nelson, McEvoy, & Schreiber, 1998). The easy and hard word lists did not differ significantly on this characteristic (see Appendix D). Connectivity is an index of the average associate-to-associate connectivity among the associates of the cue and of the target (Nelson, McEvoy, & Schreiber, 1998) (see Appendix D). This measure differed between the easy and hard words, with the hard words being less connected than the easy words. The final semantic index considered, concreteness, which is a measure of the ease with which a word can be imagined as measured on a scale from 1-7, did differ between easy and hard word lists (see see Appendix D ).

*Recording procedures.* The stimuli were recorded in a sound-attenuated booth in the Department of Communication Sciences and Disorders at the University of South Florida. The speakers were given the stimulus words to read over to allow for familiarization and to ensure fluent speech during recording. These words were provided on a sheet of paper, and the speakers were instructed to read the words at a normal conversational pace and to leave about two seconds between words. The speakers were

instructed to repeat a word when they made any type of mistake, such as a hesitation, mispronunciation, or dysfluency. To avoid ambiguity, the experimenter demonstrated an acceptable pace by reading a list of practice, non-stimulus words to the speaker. Once the speakers demonstrated understanding of the task and had familiarized themselves with the stimuli, recording began. The experimenter exited the booth and returned to the recording equipment to monitor the recording levels while the speakers read the words she instructed the speakers to repeat any target items if needed. During recording, the speakers first read a practice list. The practice list contained 10 items and familiarized the speakers with the task and allowed the experimenter to monitor and adjust the recording level (see Appendix E). Finally, the speakers were recorded reading the main word list. Distracter (non-target) words were added at the beginning of the list, the end of the list, and at the end of each column in order to avoid prosodic differences in pronunciation due to list beginning and end effects (see Appendix F). After reading all 150 words (6 distracter words and 144 stimuli words) once, the participants had completed the speaking task. This procedure took approximately 30 minutes to complete.

Ten practice words were recorded by an additional female, native English speaker in the same manner as described above. These words were used in the practice tasks for the listeners prior to the main experiment.

The speakers were recorded digitally at a sampling rate of 44.1 kHz, with 16-bit amplitude resolution, using a digital audio workstation (Roland VS890HD) and a high-quality microphone (Audio-Technica, AT4033). The words produced by the speakers were saved to the workstation and transferred digitally to computer for subsequent digital editing. Each target word was edited from the list using acoustic editing software, saved

to a separate file, and then peak normalized to a pre-specified RMS level (approximately 20 dB less than the system maximum amplitude).

*Noise mixing.* Pilot testing was conducted in order to determine the noise level at which a relatively equal challenge would be present for each of the listening groups in the main experiment. This was done by looking at the percent correct responses for each listener group in the pilot study at various signal-to-noise ratios (SNR) and estimating the SNR at which there was a 25% reduction in performance compared to the quiet condition. It was predicted that the LL group would need less noise to achieve the same decrement in performance relative to the EL and monolingual groups.

As with the main experiment, two groups of listeners participated in the pilot study: one group of monolingual English speakers ( $n = 12$ ) and two groups of bilinguals ( $n = 12$ ) who differed according to age of immersion in an English-speaking environment. The stimuli included the words used in the main experiment, divided into six lists of 24 words. The words were spoken by a female, monolingual speaker of English.

The noise used consisted of multi-talker babble from the Speech Perception in Noise (SPIN) sentences (Bilger, Neutzel, Rabinowitz, & Rzeczkowski, 1984). To avoid any potential learning effects that might result from using the same segment of noise for all the words, a two-minute segment was selected from the SPIN sentences babble which had relatively stable levels of noise throughout. The noise was then mixed with each target word by using a computer program that first randomly chose a section of the two-minute babble that was equal to the duration of the target word plus 1000ms (500 ms lead and 500 ms lag). The program then scaled the noise to achieve the desired SNR, based

on peak amplitude of the two items (word and noise), mixed the noise and word, and then rescaled the combined file to the original RMS amplitude.

During pilot testing, each listener heard one list of words at a time presented either in quiet or mixed with noise at several SNR that decreased from 18dB to 2dB in four dB steps (e.g., 18dB, 14dB, 10dB, 6dB, and 2dB). The SNR needed to obtain 75% of performance in quiet was calculated for each group. The SNRs chosen for the groups to be used in the main experiment were as follows: monolinguals had +6 dB SNR, EL had +5 dB SNR, and the LL group had +13 dB SNR.

### *Materials*

Receptive vocabulary size was measured using the Peabody Picture Vocabulary Test –Third Edition (PPVT-R; Dunn & Dunn, 1997). The PPVT is a measure of receptive vocabulary of English as well as a screening test of verbal ability for individuals aged 2-90+ years. For each target word, spoken by the test administrator, the participant must select from among four black and white drawings. It took approximately ten to 15 minutes to administer. The PPVT correlates well with other measures of receptive language, including an average correlation of .69 with the Oral and Written Language Scale Listening Comprehension subtest. Additionally, it has even higher correlations with some measures of verbal ability such the WISC-III VIQ (.91) and the KBIT (.81) (Dunn & Dunn, 1997).

In addition, receptive language was assessed for all participants using the Listening Comprehension subtest of the Oral and Written Language Scale (OWLS; Carrow-Woolfolk, 1995). For this test, the examiner read a stimulus sentence to the participant who then indicated the correct picture from among four choices that

corresponded to the stimulus. It is designed for individuals aged three through 21 years. Items probe lexical knowledge, understanding of syntactic constructions such as embedded sentences and subordination, knowledge of supra-linguistic structures such as figurative language, and other higher-order thinking skills. The OWLS correlates well with other measures of language including the CELF-R Oral Comprehension subtest (.91). The OWLS Listening Comprehension subtest also correlates well with cognitive measures that assess both verbal (WISC-III Verbal IQ, .77; K-BIT Vocabulary subtest, .76) and non-verbal ability (WISC-III Performance IQ, .70; K-BIT Matrices subtest, .59). Its correlations with global measures of cognitive ability are .76 for the WISC-III Full Scale IQ and .72 for the K-BIT Composite (Carrow-Woolfolk, 1995). These language tests were administered to examine the relationship between vocabulary and word recognition skills and more general linguistic competence in bilinguals and monolinguals.

A detailed language background questionnaire was provided to the non-native participants for collection of data on age of acquisition, language dominance, language usage, and history of speech and hearing impairment. A similar but less detailed language background questionnaire was provided to the native speakers to ensure that they were indeed monolingual and did not have a history of speech or hearing impairment (see Appendices A and B).

#### *Procedure*

Participants in the listening task were tested individually or in groups of up to four in a speech perception lab in the Department of Communication Sciences and Disorders. Upon arrival in the lab, they were greeted, told about the nature of the study, and given informed consent materials. Participants completed all consent forms, language

background questionnaires, and a basic hearing screening prior to the main experimental task. After all consent forms, hearing screenings, and language measures were completed, participants were directed to have a seat in front of one of the computers. Each session began with one practice list of ten words (with no memory pre-load) consisting of non-target words (i.e., not from the easy and hard lists) spoken by the same speaker as for the main test stimuli.

For the main experimental task, participants heard words spoken by one of the native-English speaking females, in sets of 24, 12 easy and 12 hard in each set. Following the procedure used in Luce et al. (1983), prior to the presentation of each word sub-list, the participants saw either zero, three, or six digits displayed visually, one at a time, on a computer screen positioned directly in front of them at a distance that allowed for easy viewing. The participants were instructed to remember the digits (if any) in the same order as they were presented. Each digit, sampled without replacement from the digits one through nine, remained on the screen for two seconds. The interval between the presentations of each digit was one second. Next, the words in the sub-list were presented and participants typed in what they recognized after each word. The six sets of 24 words were counterbalanced across listening conditions (3 digit conditions and 2 noise conditions) using a Latin Square design with six conditions. For example, a listener might have the quiet condition first and start with three digits, so that their order of presentation would be: three digits quiet, three digits noise, zero digits quiet, zero digits noise, six digits quiet, and six digits noise. All tokens were presented over headphones for each listener at approximately 65 dB SPL. Each participant heard all the words with none repeated. The subsets of 24 words and noise conditions were counterbalanced

across listeners within each group so that each subset was heard under each digit condition and each noise condition an equal number of times within each group, according to the Latin square design. At the end of the sub-list presentation, the participants typed in the digits they remembered seeing. The procedure continued in this manner until the six 24-item sub-lists were completed.

The words were automatically scored by a customized computer program. The misspelling of a word did not necessitate its being counted as incorrect. Rather, after all words were scored automatically by the computer program, the experimenter went through the responses and counted correct any word that was obviously misspelled or that was an obvious typo based on the position of the letters on the keyboard.

Data for the receptive vocabulary measure were collected after the main experimental task. For the PPVT, an answer book was used that had four pictures per page, with one that corresponded to the target word. The investigator said a word and the participant was instructed to say the number or point to the picture that corresponded to the word. Next, the OWLS was administered. This test also used an answer book that had four pictures per page, with one corresponding to the correct answer. For this test, however, a short sentence or paragraph was read. Participants were instructed to point to the picture or say the number of the picture that corresponded to the situation described.

Following the language testing, the 144 words presented in the main experimental task were presented again in random order over headphones along with the written counterparts in the same order. In this post-test, listeners rated each word for familiarity on a scale from 1 (not known) to 9 (very familiar). After rating all 144 words for familiarity, the participants were compensated for their participation and dismissed.

## Results

The specific research objectives for this study were (1) to compare the effect of phonological neighborhood characteristics on word recognition between native-speaking listeners and non-native listeners, (2) to examine the effect of non-native listeners' age of immersion and vocabulary level on recognition of words and recall of digits in noise and in quiet, (3) to compare the effects of increasing number of digits to recall on the recognition of easy and hard words for native and non-native listeners in both quiet and noise, and (4) to explore the effects of semantic characteristics of words on their recognition.

It was predicted that the easy/hard word effect on recognition would be greater for non-natives, especially in noise and with increased number of digits to recall. It was also predicted that vocabulary level would be the best predictor of hard word recognition compared to age-of-immersion and length-of-residence and that positive correlations would exist between the receptive vocabulary level of the non-native listeners and their word recognition scores. The effects of age of immersion and noise were also predicted to have an effect on word recognition and the recall of digits such that the later-learning (LL) group would experience more deleterious effects of noise on their ability to recognize the words and recall the digits than would the early-learning (EL) group. The possible effects of the semantic characteristics of the words on word recognition was explored.

In order to answer some of these questions, a four-way analysis of variance was conducted to analyze the percent-correct scores for word recognition task. The dependent variables were percent of words recognized correctly. Listener group (three levels: native (MO), EL non-native, and LL non-native) was the between-subjects variable; digit recall condition (three levels: 0, 3 or 6 digits), word type (two levels: easy and hard), and noise condition (quiet and noise) were the within-subjects variables. Data were converted to Rationalized Arcsine Units (RAU) (Studebaker, 1985). Doing linear tests on proportional data can be difficult since the distributions of these values are not strictly Gaussian, especially when the proportions are near 0 or 1. The Rationalized Arcsine Transform linearizes the proportions and converts them to rational arcsine units so that linear tests can be performed on the RAU values. Arcsine transformations have been used in research to transform proportions to make them more appropriate for statistical analysis, but the arcsines did not always show a clear relationship to the original proportions, making them difficult to interpret. The RAU Transform, on the other hand, produces values that are numerically close to the original percentage values over most of the range while retaining all of the desirable statistical properties of the arcsine transform.

#### *Word Recognition*

There were significant main effects of group, noise condition, and word type but not digit condition, see Table 4 and 5 for means. The MO group correctly recognized more words than did the EL group and the EL group correctly recognized more words than did the LL group ( $F(2, 93) = 32.603, p < .000, \eta_p^2 = .412$ ). Further, all groups performed better in quiet than in noise, ( $F(1, 93) = 888.16, p < .000, \eta_p^2 = .905$ ). Finally,

all groups had better word recognition for easy words than for hard words ( $F(1,93) = 299.23, p < .000, \eta_p^2 = .763$ ).

Table 4. Means of Word Recognition for Easy and Hard Words by Listener Group.

	Monolingual				Early Learners				Later-Learners			
	Easy Qt	Easy Ns	Hard Qt	Hard Ns	Easy Qt	Easy Ns	Hard Qt	Hard Ns	Easy Qt	Easy Ns	Hard Qt	Hard Ns
Means	97.4	74.5	92.7	62.0	94.9	73.1	88.7	58.1	84.3	73.1	55.0	45.5
SEM	2.65	8.95	5.48	9.32	4.58	8.96	7.97	13.29	11.01	13.10	23.25	21.82
SD	.44	1.49	.91	1.55	.76	1.48	1.33	2.21	2.25	2.67	4.75	4.45

Table 5. Means of Digit Recall by Listener Group.

	Monolingual				Early Learners				Later-Learners			
	3 Qt	3 Ns	6 Qt	6 Ns	3 Qt	3 Ns	6 Qt	6 Ns	3 Qt	3 Ns	6 Qt	6 Ns
Means	2.81	2.78	4.81	4.61	2.81	2.72	4.97	4.61	2.78	2.46	4.5	4.69
SEM	.10	.11	.25	.29	.13	.12	.21	.30	.08	.23	.40	.34
SD	.62	.64	1.47	1.74	.76	.70	1.25	1.77	.41	1.1	1.7	1.66

The interactions addressed the research questions concerning the effect of age of immersion of the non-native listeners on recognition of words in quiet and noise and the effect of phonological neighborhood characteristics on word recognition for listeners in quiet and noise. The analysis showed differences among groups as a function of the noise condition ( $F(2, 93) = 48.73, p < .000, \eta_p^2 = .512$ ), see Figures 1 and 2. The post hoc Bonferroni analysis revealed that the MO did not differ from the EL group. However, they both recognized more words than did LL and these differences were greater in quiet than noise. There was also an interaction of word type by group ( $F(2, 93) = 31.37, p < .000, \eta_p^2 = .403$ ) such that the LL recognized fewer words than the MO and the EL and

this difference was greater for the hard words than for the easy words.

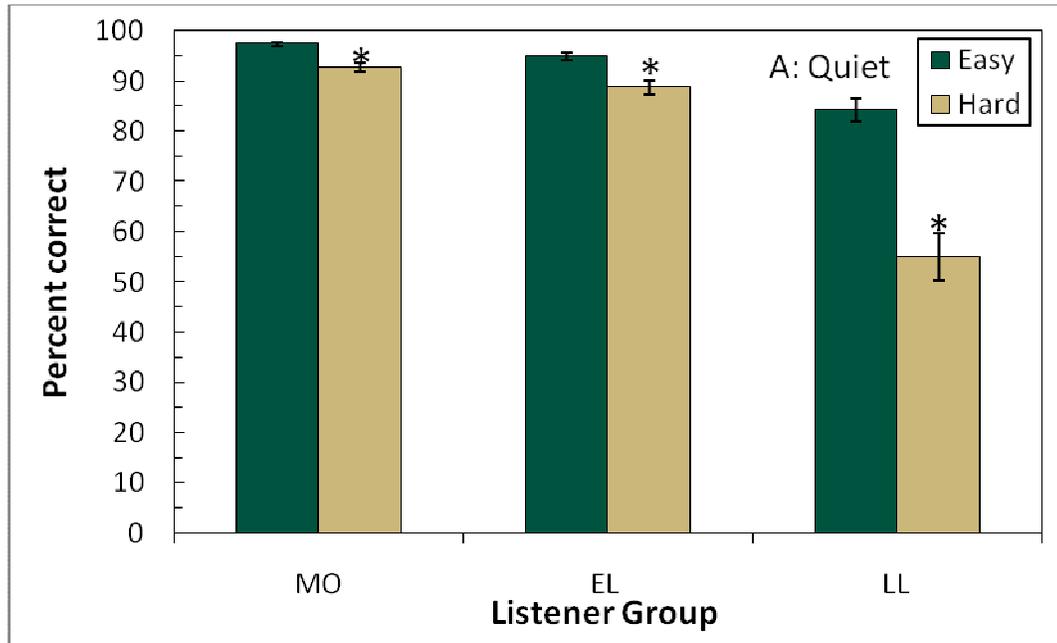


Figure 1. Percent correct word recognition for easy and hard words for all listener groups in quiet. (Error bars represent one standard error of the mean).

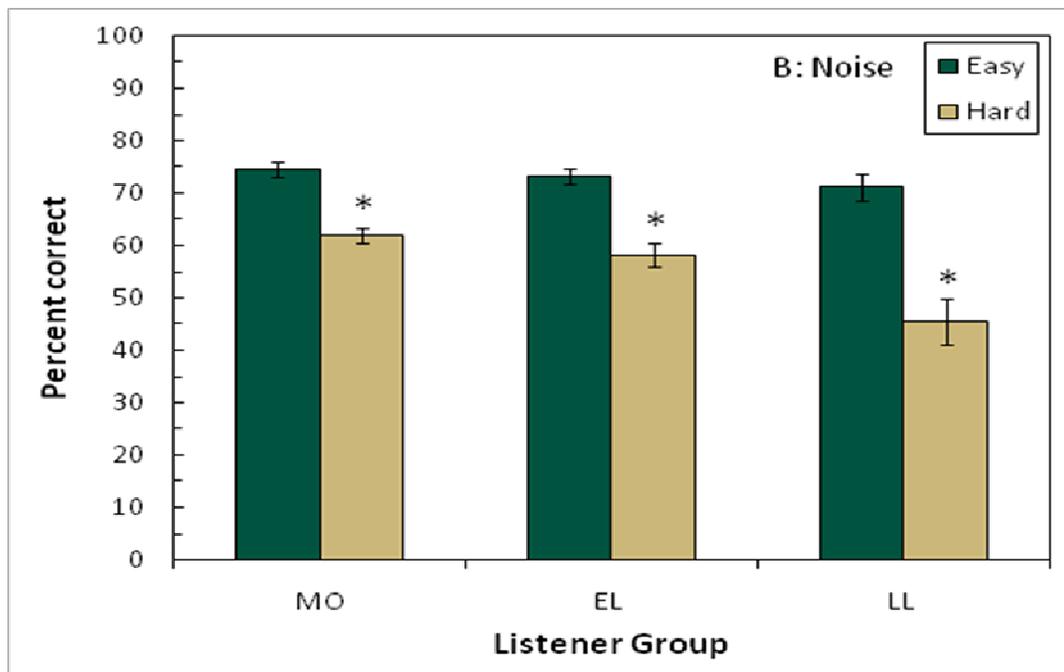


Figure 2. Percent correct word recognition for easy and hard words for all listener groups in noise. (Error bars represent one standard error of the mean).

Another objective of the present study was to determine if there was an easy-hard word effect that differed by group and whether this effect was moderated by the addition of noise. It was predicted that there would be an interaction of age of immersion with word difficulty such that the LL group would show an even greater easy/hard word effect than EL group, especially in conditions of noise. The hypothesis was confirmed by a significant three-way interaction of noise, word type, and group ( $F(2, 93) = 3.32, p < .040, \eta_p^2 = .07$ ). An examination of the paired contrasts showed that in both quiet and noise, all three listener groups had better word recognition for easy words than for hard words. However, the difference between the easy and hard word recognition accuracy was most pronounced for the LL group. In other words, the LL group showed a greater easy-hard word effect compared to the MO or EL groups. Interestingly, the difference between easy and hard word recognition scores among the groups lessened in the noise condition (see Figures 1 and 2).

To further explore the question regarding the effects of phonological neighborhood on word recognition, analyses were conducted to determine the frequency with which participants chose a neighbor versus a non-neighbor when they incorrectly identified the target word. The motivation for these analyses was the need to gain some insight into the phonological neighborhoods of non-native listeners compared to native listeners. The Neighborhood Activation Model (Luce & Pisoni, 1998) predicts that low-frequency words with many and more frequent neighbors will be more difficult to recognize because of competition from the neighbors. This supposes that, when in error, a listener is likely to choose a neighbor, especially for the hard words. I wanted to know whether this was as true for non-native listeners as it is for native listeners, and to see to

what extent vocabulary level moderates this process. Two sets of analyses were conducted: one compared the groups in the rates at which participants chose neighbors, and the other compared the word type with regard to the rates at which a neighbor was chosen. For the first, a one-way analysis of variance, with listener group (three levels: MO, EL, and LL) the between-subjects variable, was performed on the percent of neighbors chosen when a target was not accurately recognized. “Neighbor” was operationalized as a word that differed from the target by one phoneme, such as “rat” for “cat.” Non-word neighbors, such as “dat” for “cat,” were not counted as neighbors. There was a significant difference among the groups in the percentage of neighbors chosen for the easy words ( $F(2, 93) = 6.693, p < .002$ ) and the hard words ( $F(2, 93) = 12.30, p < .000$ ). Tukey Honestly Significant Difference (HSD) post hoc analysis revealed that, for the easy words, the MO and the EL groups did not differ from each other, but they both chose more neighbors than non-neighbors relative to the LL group, see Table 6 for means. For the hard words, the MO chose neighbors significantly more than did the EL group, who chose more neighbors than did the LL group.

Table 6. Mean percentage of times groups chose a neighbor when in error during word recognition task.

	Monolingual		Early Learners		Later-Learners	
	Easy	Hard	Easy	Hard	Easy	Hard
Means	57.89	84.82	57.02	78.91	44.19	71.48
SEM	2.62	1.52	2.40	1.80	3.35	2.23
SD	15.73	9.13	14.40	10.78	16.39	10.90

A paired samples t-test was conducted to determine whether a difference existed in the rate at which a neighbor versus a non-neighbor was chosen between the easy and hard words for each of the groups overall. It was found that participants were more likely

to choose a neighbor in the hard condition than the easy condition. This makes sense considering that there are fewer neighbors from which to choose for the easy words, and the target tends to be more frequent than those neighbors. In contrast, the hard-word neighborhoods have many more neighbors from which to choose, and those neighbors tend to be more frequent. Finally, an analysis was conducted to determine whether the difference in likelihood to choose a neighbor between the easy and hard words varied among groups. This was done by calculating the difference in the percentage of times a neighbor was chosen when in error between the easy and hard words and then comparing this difference among the groups using a one-way analysis of variance. This difference was not significant. Anecdotally, the LL group was observed to choose neighbors that are not English words (and possibly non-words in their L1 also), something that the MO and EL groups did not do.

#### *Vocabulary, Language Scores, and Familiarity Ratings*

It is argued that the ability to discriminate and subsequently encode hard words is a skill that develops with knowledge of the language (i.e., vocabulary level), so, it was predicted that the EL listeners in this study would have better recognition of the words than the LL listeners. It was further hypothesized that vocabulary level would be the best predictor of hard word recognition. To investigate the effect of vocabulary level on recognition of words, a set of one-way analyses of variance was conducted comparing the groups' vocabulary scores, language scores, and familiarity ratings. All three groups differed significantly from each other in their vocabulary scores, with the MO group scoring the highest, followed by the EL bilinguals, and then the LL bilinguals ( $F(2, 93) = 26.54, p < .000$ ), as shown in Figure 3. The groups also differed significantly from one

another in their listening comprehension scores, again with the MO group scoring the highest, followed by the EL bilinguals and the LL bilinguals ( $F(2,93) = 21.87, p=.0001$ ).

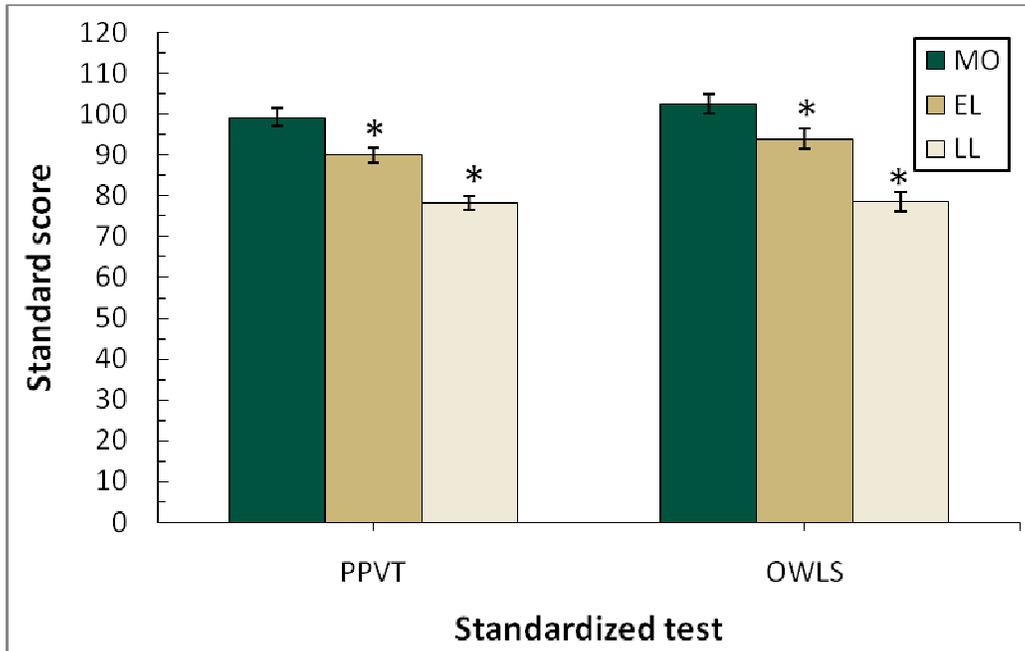


Figure 3. PPVT and OWLS standardized scores for all listener groups.

A familiarity rating of 5.5 or higher on the nine-point scale was given by the MO group for 143 of the 144 words, 138 of the words by the EL group, and 118 of the words by the LL group. All but two of the lower-rated items were hard words. A one-way analysis of variance was conducted to determine whether differences existed among the groups in how they rated the words for familiarity. Though all groups rated the easy words as more familiar than the hard words, there were group differences in how they rated these words for familiarity, ( $F(2, 93) = 4.304, p = .016$ ) (see Table 7 for means). Post hoc LSD tests revealed that the MO and the EL groups did not differ significantly from one another in how they rated the easy words. However, the LL group rated the easy words as significantly less familiar than both the MO and EL groups. For the hard words,

the same pattern of results was found. That is, the MO and EL groups did not differ from each other in how they rated the hard words for familiarity, but both groups differed significantly from the LL ( $F(2,93) = 29.56, p < .001$ ).

Table 7. Vocabulary and Language Tests, Familiarity Ratings, and Demographic Variables

	Monolingual				Early Learners				Later-Learners			
	PPVT	OWLS	FE*	FH**	PPVT	OWLS	FE*	FH**	PPVT	OWLS	FE*	FH**
Means	99.22	102.44	8.80	8.21	89.94	93.94	8.85	7.87	78.17	78.58	8.64	6.31
SEM	2.07	2.35	.05	.161	1.82	2.45	.02	.12	1.73	2.34	.06	.26
SD	12.43	14.07	.33	.97	10.90	14.70	.13	.70	8.46	11.47	.32	1.27
Range	55	51	1.22	4.40	48	52	.47	2.78	36	48	1.35	5.15
Min	72	77	7.78	4.6	72	73	8.53	6.14	60	53	7.65	3.39
Max	127	128	9	9	120	125	9	8.92	96	101	9	8.54

\*FE= average familiarity rating for the easy words

\*\*FH= average familiarity rating for the hard words

*Correlational Analyses.* In order to further explore the relationship between the word recognition scores and various demographic factors, vocabulary, and language skills, a series of correlational analyses (Pearson's  $r$ ) was performed. It was predicted that vocabulary level would be the best predictor of hard word recognition and that positive correlations would exist between the vocabulary level of the non-native listeners and their word recognition scores. Thus, it was necessary to look at the correlations of age-of-immersion (AOI), length-of-residence (LOR), vocabulary level, and receptive language scores with word recognition in order to determine how strongly correlated each of these variables was with word recognition accuracy. Then, regression analyses were conducted to find which variables predicted word recognition accuracy.

In each case, a correlation was run with all the groups together and in some cases, separate correlations were run for each group. The correlations among the different variables for all groups considered together can be seen in can be seen in Table 4.

Table 8. Correlations among overall word recognition, vocabulary level (PPVT), receptive language scores (OWLS), and demographic variables (AOI and LOR) for all listeners.

	<i>Average Word Recognition</i>	<i>PPVT</i>	<i>OWLS</i>	<i>AOI</i>	<i>LOR</i>
Average Word Recognition		.601*	.572*	-.702*	.478*
PPVT			.548*	-.526*	.396*
OWLS				-.460*	n.s.
AOI					-.656*
LOR					

\* $p < .005$ , (2-tailed)

As shown in Table 8, AOI was significantly negatively correlated with overall word recognition scores (collapsed across easy and hard words in both quiet and noise) ( $r = -.702$ ,  $p < .001$ ) for the non-native groups (note that the MO group did not have values for AOI, as with LOR, so this statistic reflects that of the non-native groups only). PPVT scores were also significantly correlated with overall word recognition for the samples as a whole ( $r = .601$ ,  $p < .001$ ). Receptive language scores, as measured by the OWLS, were also significantly correlated with overall word recognition scores ( $r = .572$ ,  $p < .001$ ). Finally, LOR was significantly correlated with overall word recognition scores for the bilingual groups ( $r = .478$ ,  $p < .001$ ).

Tables 9 and 10 show the correlations among the word recognition scores in the various conditions (easy, hard, quiet, and noise) with vocabulary level, receptive language scores, and familiarity ratings for the non-native groups. Word recognition scores were not significantly correlated with any variables for the MO group.

Table 9. Correlations among spoken word recognition, vocabulary, listening comprehension scores, word familiarity, and demographic variables for EL group

	<i>Easy Word Recognition</i>	<i>Hard Word Recognition</i>	<i>Word Recognition in Quiet</i>	<i>Word recognition in Noise</i>
PPVT	.425**	.339*	.572**	n.s.
OWLS	n.s.	.412**	.452**	n.s.
Familiarity Easy Words	n.s.		n.s.	n.s.
Familiarity Hard Words		n.s.	.324*	n.s.
AOI	-.335*	-.281*	n.s.	-.320*
LOR	.352*	n.s.	n.s.	n.s.

\*  $p < .05$  (2-tailed) \*\*  $p < .005$

Table 10. Correlations among spoken word recognition, vocabulary, listening comprehension scores, word familiarity, and demographic variables for LL group

	<i>Easy Word Recognition</i>	<i>Hard Word Recognition</i>	<i>Word Recognition in Quiet</i>	<i>Word recognition in Noise</i>
PPVT	.499**	.606**	.565**	.579**
OWLS	.485**	.464*	.527**	.424*
Familiarity Easy Words	.559**		.723**	.659**
Familiarity Hard Words		.411*	.391*	.355*
AOI	-.531**	-.661**	-.582**	-.658**
LOR	n.s.	n.s.	n.s.	n.s.

\*  $p < .05$  (2-tailed)

\*\*  $p < .005$

Word recognition scores for the two non-native groups, however, were significantly correlated with many of these variables. For the EL, vocabulary scores, LOR, and AOI were moderately correlated with easy-word recognition. The same was true for the LL group, except that their easy-word recognition scores were also correlated with their familiarity ratings of the easy words and not with LOR.

For the hard words, the EL and the LL groups' vocabulary level, receptive language skills, and AOI were moderately correlated with word recognition. However, for the LL group, familiarity ratings were also correlated with hard-word recognition.

In looking at all the words in the quiet condition, the EL and LL groups' vocabulary level, receptive language scores, and hard-word familiarity ratings were significantly correlated with word recognition. The LL group's word recognition in quiet was also significantly correlated with easy-word familiarity ratings and AOI.

In the noise condition, AOI was significantly correlated with word recognition for the EL group. However, the LL group's word recognition in noise was significantly correlated with AOI, vocabulary and receptive language scores, and familiarity ratings.

Next, a correlation was conducted that investigated the relationship among the size of the easy-hard word effect (calculated as the difference between the easy and the hard word scores) and vocabulary level, receptive language, and demographic variables. As noted, the size of the easy-hard word effect was found to be greatest for the LL group relative to the NS and EL groups. Vocabulary and language scores as well as the demographic variables were all significantly correlated with the size of the easy-hard word effect for all groups: PPVT ( $r = -.464, p < .001$ ), AOI ( $r = .690, p < .001$ ), LOR ( $r = -.373, p = .003$ ), and OWLS ( $r = -.496, p < .001$ ). In other words, as vocabulary level,

receptive language skills, and length of residence increase, the size of the easy-hard word effect decreases. However, as age-of-immersion increases, (i.e., the later listeners were immersed in an English-speaking environment) the size of the easy-hard word effect gets larger.

*Regression Analyses.* It was expected that vocabulary level of the non-native speakers would predict their word recognition accuracy. In order to confirm this prediction, stepwise multiple regression analyses were conducted in order to find the variables most predictive of performance on the word recognition task. An analysis was first performed which collapsed across group, word type, and noise condition. In other words, it looked for the predictive variable for the overall word recognition scores for all listeners. The independent variables were PPVT scores (vocabulary), OWLS scores (receptive language), and word familiarity ratings for the easy and hard words. For this analysis, the variables AOI and LOR could not be entered into the equation because the NS group did not have values for them. PPVT accounted for 35.5% of the variance ( $\beta = .601, p < .001$ ), with OWLS ( $\beta = .347, p < .001$ ) contributing an additional 8.4% of the variance, and easy-word familiarity ratings ( $\beta = .262, p = .001$ ) contributing an additional 6.3%.

The next set of analyses considered the non-native speaker group only. Again, four stepwise multiple regression analyses were performed with word recognition in quiet, noise, easy words, or hard words as the dependent variables, as described below. The independent variables were PPVT scores (vocabulary level), OWLS scores (receptive language), age-of-immersion, length of residence, and word familiarity. For the EL group, PPVT was found to be a significant predictor of word recognition for the

easy words ( $\beta = .425, p = .010$ ) contributing 15.6% of the variance, with LOR ( $\beta = .363, p = .017$ ) contributing an additional 13.2% of the variance. For the LL group, word familiarity was found to be a significant predictor of easy word recognition ( $\beta = .559, p = .005$ ) contributing 28.1% of the variance.

For the hard words, the EL group's OWLS scores were found to be a significant predictor of hard word recognition ( $\beta = .412, p = .013$ ), contributing 14.5% of the variance. For the LL group, AOI ( $\beta = -.504, p = .003$ ) contributed 41.2% of the variance and, PPVT scores ( $\beta = .416, p = .012$ ) contributed an additional 14.8% of the variance.

In looking at all the words in the quiet condition, the EL group's PPVT scores ( $\beta = .572, p < .001$ ) accounted for 30.7% of the variance, with receptive language ( $\beta = .323, p = .024$ ) accounting for an additional 9.7% of the variance. For the EL group, 50.2% of the variance was accounted for by their familiarity with the easy words ( $\beta = .583, p = .001$ ), and their AOI accounted for an additional 9.7% of the variance ( $\beta = -.342, p = .013$ ).

For the noise condition, the EL group's vocabulary level accounted for 10.6% of the variance in the word recognition scores ( $\beta = .362, p = .030$ ). For the LL group, two variables were found to be significant predictors of easy word recognition, familiarity ( $\beta = .468, p = .005$ ) and AOI ( $\beta = -.465, p = .005$ ), accounting for 40.9% and 17.9% of the variance, respectively.

### *Digit Recall*

Another set of research questions concerned the effects of age-of-immersion, word-type, and noise on the recall of digits using the same memory-preload technique as Luce and Pisoni (1983). This task consisted of presenting the subjects with a list of digits

that they were to rehearse throughout the primary word-recognition task. It was predicted that the LL group would experience more deleterious effects of noise on their ability to recall the digits than would the EL group. A three-way analysis of variance was conducted to analyze the number of digits correctly recalled and the percent correct words recognized in the various digit conditions. The dependent variables were the number of digits recalled correctly and the percentage of words accurately recognized. Listener group (three levels: MO, EL, and LL) was a between-subjects variable. Noise condition (two levels: quiet and noise) and digit recall condition (three levels: 0, 3, or 6 digits) were within-subjects variables. The prediction was not supported because there were no significant effects found for the recall of digits in this study.

#### *Semantic Features*

There were no predictions regarding the effects of the semantic characteristics of the words on word recognition because I did not have data for all the words nor did I control for the factors. Instead, the analyses were exploratory in nature with no a-priori expectations. Correlational analyses were conducted in order to investigate the relationship between the semantic features of the words and the accuracy with which they were recognized. There were no significant correlations found between the word recognition scores and any of the semantic characteristics for the subset of stimuli for which semantic characteristics could be calculated.

## Discussion

This study assessed the contributions of phonological neighborhood characteristics and memory load on spoken word recognition by monolingual English listeners and two groups of non-native bilingual listeners who differed in their age of immersion in an English-speaking environment. The prediction that the easy/hard word effect on recognition would be greater for non-natives was supported. The prediction that earlier-learning listeners would have better recognition of the words compared to the later-learning listeners was confirmed and supports the hypothesis that the ability to discriminate hard words is a skill that develops with knowledge of the language. Further, the prediction that vocabulary level would correlate with word recognition accuracy such that those with lower vocabulary scores would also do more poorly on the word recognition task was also supported. The predictions regarding the effects of the recall of digits on word recognition and the effects of phonological neighborhood, noise, and age-of-immersion (AOI) on the recall of digits were not supported, which may have been due to a failure to manipulate memory load adequately. Finally, although the semantic characteristics of the word were not controlled and no predictions regarding their effects were made, they were nevertheless explored. No significant effects of the semantic characteristics of the words on the accuracy with which they were recognized were found.

### *Phonological Neighborhood*

The stimuli in this study consisted of the easy and hard words used by Bradlow and Pisoni (1999). In their study, easy words had higher intelligibility than hard words, and this was especially true for the non-native listeners. They argued that the ability to make the fine acoustic-phonetic distinctions required to discriminate the hard words is a skill that develops with knowledge of the sound-based system of the language. They found that all listeners identified words that were more easily discriminated from other words in their neighborhoods compared to words with many similar sounding neighbors, and this easy/ hard word effect was greatest for the bilingual listeners.

In the present study, a substantial easy-hard word effect was obtained only for the earlier-learner listeners. Furthermore, oral vocabulary size was significantly correlated with performance for the non-native listener groups only. Thus, the greater easy-hard effect for non-native listeners can be explained as an effect of both phonetic proficiency and vocabulary size on the structure of lexical neighborhoods, and it seems that these skills are integrated.

Garlock et al. (2001) noted that high-frequency words are more likely to overlap with many other words on a segmental basis, and the words with which they overlap also tend to be high in frequency. The implications for this study are that the non-natives, especially the later-learners, may have different neighborhoods than the monolinguals, so that the words in their neighborhoods would overlap less (i.e., they have fewer English words in their neighborhoods, so less overlap is likely to result). However, these same English words might also overlap with words in their native language, causing greater difficulty with word recognition tasks. Later-learners may be likened to language-

learning children. Garlock et al. (2001) argued that children show a smaller competition effect than adults because they do not know as many words, so there are fewer competitors in their neighborhoods. They maintained that children's representations may not be as differentiated as adults', so words from dense neighborhoods are not as impeded relative to words from sparse neighborhoods.

Alternatively, because children's representations are undergoing significant change, the recognition of hard words might be especially difficult and the recognition should be best for words that are most familiar and likely to be more stable and robust. This hypothesis could be extended to non-native speakers, especially for later-learners with relatively small vocabulary sizes. It seems reasonable to assume that, like children learning a first language, later-learners' representations are also undergoing significant change. Following this argument, their representations for easy words might be more established than those of the hard words. If this were the case, performance would be expected to be more similar between the earlier-learners and later-learners for the easy words compared to the hard words. The later-learners might also have less competition for accurate word recognition because they do not have as many words in their neighborhoods. As Marian and Blumfield (2006) explained, in a native language, word recognition tends to be better for words that are used often. In a non-native language, however, a L2 learner may have more limited exposure to and use of particular high-frequency words. This might give those words the same status as low-frequency words, effectively making the neighborhood effects more pronounced in the L2.

If later-learners have sparser neighborhoods overall, then the density of their neighborhoods might be expected to increase as their L2 vocabularies increase. The

increased vocabulary (and by extension, proficiency) in the L2 would then allow greater ability to make fine-grained phonological distinctions between neighbors (Imai et al., 2005). As evidenced by their lower PPVT scores, the non-native listeners in this study had smaller L2 vocabularies, and thus, target words may have had fewer competitors. The results of this study support this line of reasoning: the easy-hard word effect was larger for the later-learner group compared to the earlier-learner and monolingual groups.

Luce and Pisoni (1998) argued that the relatively sparse phonological neighborhoods of children cannot be explained by smaller vocabularies alone. They suggest that, because of the small size of their neighborhoods, children may use recognition strategies that are more holistic rather than segmental. This is because the fine-grained phonetic discrimination strategies adults used are not necessary given that children's neighborhoods are not as densely populated. The description of the developmental path of phonological neighborhood effects could shed light on the processes involved in word recognition for non-natives. Non-natives, especially those who are later learners, certainly have smaller L2 vocabularies than native speakers. However, their neighborhoods may include words from their L1, which could compete with L2 words. The degree of phonological overlap between the first and second language may impact bilingual word recognition in the L2. In fact, Boukrina and Marian (2006) manipulated cross-linguistic phonological overlap between Russian and English in a lexical decision task and found that, as phonological overlap increased, so did the speed and accuracy of responses in the L2, but not the L1. They suggested that facilitation of L2 lexical decisions occurred because of co-activation of wider L2 phonetic categories with similar L1 sounds.

The degree of activation of the target (in this case, English) versus the non-target language (the listeners' various L1s) could vary with the level of proficiency of the individual. This was evident in the present study by the rate at which the groups chose neighbors versus non-neighbors when they incorrectly identified the target word. All three groups were more likely to choose a neighbor for the hard words than the easy words. This supports the findings of Roodenrys et al. (2002) who found that neighbors were more likely to be chosen if the target was less frequent and had many high-frequency neighbors. However, non-natives, especially the later-learner group, were more likely to choose a non-neighbor when they had incorrect word recognition. For the easy words, when in error, the monolingual and earlier-learner group chose a neighbor 57% of the time, while the later-learner group chose a neighbor 44% of the time. For the hard words, the monolingual group chose a neighbor 84% of the time when in error, the earlier-learner group chose a neighbor 78% of the time, and the later-learner group chose a neighbor 71% of the time.

Support for the Neighborhood Activation Model (Luce & Pisoni, 1998) comes from the finding that neighbors were more likely to be chosen over non-neighbors for the hard words relative to the easy words. Easy words have fewer neighbors competing during recognition tasks, whereas hard words have many more neighbors from which to choose that are higher in frequency compared to the target word. What is interesting is that the earlier-learner group chose neighbors at roughly the same rate as the monolinguals for the easy words, and the later-learner group chose significantly fewer neighbors than the other groups. For the hard words, however, the later-learner group chose significantly fewer neighbors compared to the other groups and the earlier-learner

group chose neighbors at a rate somewhat in between that of the native speakers and the later-learner group. So, the rate at which one chooses a neighbor when in error seems to increase with vocabulary size in the L2, especially for the easy words. The fact that the later-learner group chose more non-word neighbors than the other groups is also interesting. Perhaps this group was less aware that some of their responses were non-words.

It has been suggested that the degree of confusability for given words conveys information about the listener's internal lexicon and the relative accessibility of its component words (Goldinger et al., 1991). For this reason, receptive vocabulary size and listening comprehension scores were gathered to provide insight into their internal lexicons.

#### *Proficiency and Vocabulary Level*

Proficiency level of non-native speakers has not been considered in studies that have investigated the easy/ hard word effect and cognitive demand using the digit recall technique. Spoken word recognition by non-native speakers depends largely on vocabulary development in the target language. Although studies exist that have considered vocabulary level of children and adults in word recognition tasks (Garlock et al., 2001), no studies have considered vocabulary level as an index of L2 proficiency for non-natives in word-recognition tasks between easy and hard words. Imai et al. (2005) correlated proficiency (defined as the degree of accentedness of the non-native speakers as measured by native listeners) with number of years of English-language study and word recognition accuracy. Likewise, Bradlow and Pisoni (1999) performed correlational analyses of factors such as age of English study onset, number of years of

English study, and number of years in an English-speaking environment, but neither study directly measured the vocabulary level of their non-native participants.

The current study used several measures to classify non-native participants as higher or lower-proficiency. The PPVT was used to obtain an objective measure of participants' vocabulary level in English, the OWLS gave an index of their oral language comprehension, and the language background questionnaire provided information regarding their language use and dominance. Whereas the native English speakers scored higher than the non-natives on all language measures, the earlier-learners scored higher on the PPVT and the OWLS and reported more frequent use and better command of English than did the later-learners. It was predicted that the vocabulary level of the bilinguals would influence their word recognition, with the earlier-learners identifying more words correctly than the later-learners. Results confirmed this prediction: monolinguals recognized more words than both bilingual groups, and the earlier-learners recognized more words than the later-learners. This finding echoes the results of previous studies that demonstrated that later-learners have more difficulty in English-word recognition tasks than do later-learners (Imai et al., 2005; Rogers et al., 2004).

The non-native groups in this study differed in the extent to which their vocabulary levels were predictive of their word recognition scores. The PPVT was the best predictor of easy word recognition for the earlier-learner group but familiarity with the easy words was the best predictor for the later-learner group. The OWLS, which is a measure of receptive language skills, was the best predictor of hard word recognition for the earlier-learners and age-of-immersion was the best predictor for the later-learners (though their vocabulary level accounted for additional variance for the latter group).

Moreover, vocabulary was strongly correlated with easy- and hard-word recognition for the earlier-learners, compared to age-of-immersion or length-of-residence. For the later-learner group, vocabulary and age-of-immersion shared similar correlations with word recognition for the hard words. For the easy words, age-of-immersion and vocabulary level were similarly correlated with word recognition scores for this group. One explanation for these findings is that the earlier-learners all had very low ages-of-immersion with a relatively restricted range compared to the later-learners, suggesting that if age-of-immersion is held comparatively constant, it is vocabulary that differentiates performance on word recognition tasks.

Marian and Blumenfeld (2006) also used the PPVT as a measure of language proficiency in their study exploring the role of phonological density in lexical access in native and non-native languages. They found it to be a better predictor of naming accuracy than age of acquisition and suggested that bilinguals can improve L2 performance with increased proficiency, regardless of age of acquisition. Their findings and those of the present study support the argument that increasing vocabulary level may result in greater attention to fine phonetic detail. Positive correlations were found between vocabulary level and word recognition scores, suggesting that lexical development influences phonological knowledge of the target language or, at least, that lower-level, phonological knowledge and higher level, lexical knowledge influence one another.

The relationship between vocabulary level and word recognition accuracy for the bilingual groups can be better understood when one considers the literature on children's spoken word recognition abilities. Garlock et al. (2001) offered two proposals regarding

the neighborhood competition effects for children as they relate to vocabulary level. On the one hand, the effects of competition might be smaller because they do not know as many words as adults; therefore, the recognition of dense (hard) words relative to sparse (easy) words might not be as difficult for them compared to adults. On the other hand, when their phonological representations are undergoing significant change, children might experience greater effects from competition, making the hard words effectively harder. The authors argued that, in order to distinguish among increasing numbers of items in the mental lexicon, spoken word representations must become more segmentally structured. Thus, it is vocabulary growth that drives changes in the lexical representation of words.

The applicability of Garlock et al.'s (2001) hypothesis to non-native speakers is apparent. The authors state that words that are least robust and stable in terms of their familiarity and neighborhood status should undergo the greatest developmental change in spoken word recognition. Hard words require fine-grained representations for accurate recognition. If differences in vocabulary development drive differences in performance between children and adults, then by extension, differences in L2 proficiency may drive differences in word recognition accuracy among non-native speakers. As with children and adults, non-native speakers who differ in L2 proficiency should demonstrate more similar word recognition for easy words because they can be recognized on a more segmental basis due to their more robust representation and less need for fine-grained distinctions. The segmental recognition of easy words is less hampered than that of hard words because they contrast with fewer words on a single phoneme basis.

### *Noise*

In this study, listeners identified English words spoken by two native English speakers and later rated the words for their familiarity. The word lists were presented in either quiet or in noise as it is expected that noise would make phonetic discrimination harder. The level of noise added to the stimuli was selected based on pilot testing and designed to cause relatively equal decrements to word recognition. Specifically, the noise level was intended to bring word recognition scores down to approximately 70% and 75% of the word recognition scores in quiet. This was mostly true for all groups: the groups' easy word scores in noise averaged between 77% and 82% of that in quiet and the hard word scores in noise averaged between 67% and 72% of that in quiet. Thus, the same level of noise had a different impact on the easy and hard words for all groups such that the hard words were more deleteriously affected, possibly because accurate recognition of hard words requires the ability to make fine phonetic distinctions among phonetic cues, some of which might be masked in noise.

### *Digit Recall*

It was predicted that the later-learners would experience more deleterious effects of noise on their ability to recall digits than the earlier-learners because of the additional short term memory capacity needed to rehearse the digits, which would subsequently leave less capacity for the encoding of words in the recognition task. Groups showed no effect on the number of digits recalled under any conditions. That is, presumably increasing the cognitive demand through the manipulation of noise, word type, or number of digits to recall did not affect recall of the digits. This may have been because the manipulation of memory load failed. Anecdotally, the participants in this study seemed

to not pay much attention to the digits during their presentation and were guessing during the recall task. It is, therefore hard to say whether cognitive demand was actually manipulated.

Luce and Pisoni (1983) were looking at the recall of synthetic versus natural words when using the memory preload technique. In other words, they used a memory task to interfere with a subsequent memory task. It was used in this study to increase cognitive demand in a recognition task. Further, they did not find differential effects of digit preload across the natural and synthetic lists. Luce and Pisoni (1983) did find that the number of subjects who recalled all the digits accurately decreased in the synthetic condition compared to the natural condition, especially for the six-digit condition relative to the three-digit condition. By extension, it was expected in this study that in the noise condition, subjects would recall fewer digits. That was not the case.

The reasons for the null findings for digit recall might become clearer when considering the model of memory by Baddeley and Hitch (1974). Baddeley and Hitch (1974) proposed the idea that memory is composed of three main components: the central executive which controls the flow of information to and from its slave systems: the phonological loop, and the visuo-spatial sketch pad. The latter two systems are short-term storage systems for the verbal and visuo-spatial domains respectively. In 2000, Baddeley added a fourth system to his model, the episodic buffer, which links information across domains with time sequencing and has associations with long-term memory and semantic meaning. The phonological loop deals with sound or phonological information and consists of two parts: the short-term phonological store which rapidly decays and an articulatory rehearsal component that keeps the memory traces active.

Auditory information is thought to enter into the phonological store, whereas visually presented speech is transformed into a phonological code by silent articulation and thereby is encoded into the phonological store. The phonological store remembers speech sounds in their temporal order, while the articulatory rehearsal component repeats the series of words to prevent them from decaying. Further, there seems to be an effect of phonological similarity such that lists of words that are similar in sound are harder to remember than words that do not sound alike. In contrast, semantic similarity does not seem to have an effect on memory, supporting the assumption that verbal information is coded phonologically in working memory.

Baddeley and Hitch (1974) found that performance of two simultaneous tasks which used two separate perceptual domains (e.g., a verbal and a visual task) is nearly as efficient as performance of the tasks individually. In contrast, performance of two simultaneous tasks requiring use of the same perceptual domain is less efficient than when performing the tasks individually. Thus, there is less interference between visual and verbal tasks than between two visual tasks or two verbal tasks. The present study required participants to remember a visually presented list of digits shown before each of the word lists was presented auditorally. Perhaps, a task in which the to-be-remembered material was an auditorally presented list of similar sounding words would have yielded an effect for the cognitive demand condition. It would also be of use to look at the reaction time data to determine if differences existed among the groups in the time it took to record the words and the digits under the various conditions. According to the model presented by Baddeley and Hitch (1974), the presentation of material through an auditory mode would have interfered more with the recognition of the auditorally presented words

because both tasks require use of the same perceptual domain. To reiterate, however, the manipulation of cognitive load most probably failed in this study.

### *Semantic Characteristics*

The null findings for semantic characteristics could be explained by the fact that many of the words did not have values and these characteristics were not manipulated in this study. Semantic network characteristics provide information about the relationships among words based on their meanings. While some of the semantic network characteristics differed between the word lists, a more controlled manipulation of these indices might reveal that the way words are semantically organized in the mental lexicon does influence their recognition. Future word recognition studies should be conducted with words in which semantic and neighborhood characteristics are varied orthogonally in order to tease apart the contributions of each. A limitation of this study is that not all the words had data points for each of the semantic features, especially the hard words. For example, only 62 of the easy words had values for connectivity, while the hard words only had 34. It should be noted that the semantic characteristics may have an effect on word recognition accuracy, but because they were not controlled in this study, it is difficult to speculate.

### *Future Directions*

This study provided valuable information about differences between native and non-native listeners in their recognition of English words and the contributions of proficiency level, neighborhood characteristics, and cognitive demand. However, future work is needed in order to more fully understand the factors that affect L2 word recognition. First, consider the stimuli. The target frequency of the words was based on

the frequency counts from the Brown Corpus of printed text<sup>1</sup> (Kucera & Frances, 1967). Though these norms are somewhat old, their use reflects the notable absence of more recent frequency counts and the relative lack of spoken word frequency counts. The implications of using frequency counts based on written texts is that the lists are very sensitive to the corpora from which they are drawn, particularly to the style, language, and content of the corpora. For example, a list generated from six million words of newspaper articles is likely to be significantly different from a list generated from six million words of internet postings or magazines. There is a newly created list of spoken word frequency counts available containing 1.6 million American-English words (Pastizzo & Carbone, 2007). This list was derived from the Michigan Corpus of Academic Spoken English (MICASE) which includes 152 transcriptions of lectures, meetings, advisement sessions, public addresses, and other educational conversations spoken by students, faculty, and other staff members and recorded at the University of Michigan. The authors found a moderately strong, positive correlation between log written frequency and log spoken frequency and suggested that a written measure can be replaced with spoken counts. Future work in second-language speech perception should, therefore, consider using spoken frequency counts.

Further consideration of the stimuli for future research could involve investigating the contributions of the semantic characteristics of words to their recognition especially for bilingual populations. Although this study attempted to explore the effects of the semantic characteristics of the stimuli, they were not manipulated or controlled in any manner. A future study could manipulate both phonological characteristics and semantic

characteristics (set size, connectivity, and concreteness) to determine if these variables affect the recognition of words.

Second, consider the types of errors that monolinguals versus bilinguals are making, specifically as they relate to neighborhoods and the changes associated with increasing proficiency level. Considering the degree of phonological and semantic (in terms of cognates) overlap between the target language and that of the non-native participants may also shed light on how different language backgrounds may affect the recognition of L2 words.

Finally, investigating the contribution of probabilistic phonotactics would provide information about how non-native listeners' mental lexicon is organized compared to that of native listeners. Phonotactics refers to a system of rules or constraints that dictate the permissibility of the occurrence of segments within syllables and words of a language (Auer & Luce, 2003). For example, in English, /la/ may legally occur at the beginning of a syllable, whereas /lda/ may not. Further, these permissible segments and their sequences occur more or less frequently in a language (e.g., /tra/ occurs frequently in English, whereas /kwa/ occurs less frequently). Probabilistic phonotactics refers to the relative frequencies of segments occurring in a listener's language (Auer & Luce, 2003).

Vitevitch et al. (1998) explained that the neighborhood density effects have a lexical focus, whereas probabilistic phonotactics effects have a sub-lexical focus. The facilitative effects of probabilistic phonotactics for non-words occur because non-words fail to activate competing lexical representations. Therefore, the processing of high-probability non-words benefits from the absence of lexical competition in the presence of high frequency segments. The easy-hard word effect could be minimized or reversed in

favor of probabilistic phonotactics by controlling the neighborhood density of the words while varying their phonotactic probability. For example, consider a word that is so unfamiliar to an individual with very low proficiency in the L2 that the word is effectively a non-word to the listener. Would recognition of the word be facilitated from its probabilistic phonotactics in ways that would not benefit a native listener because of the native listener's lexical focus? Or would the non-native require greater proficiency in the L2 before probabilistic phonotactics shows its facilitative effects?

Answers to these questions would have important implications for the teaching and assessment of non-native speakers and for the ways in which teachers can improve non-native speakers' comprehension of spoken material. Bilingual listeners have greater difficulty perceiving speech in their L2 than do native listeners, especially under adverse listening conditions and under conditions of increased cognitive load, such as noise. Thus, the findings of this study for second language pedagogy are apparent. Classrooms can be quite noisy, and the task of comprehending a lecture is made more difficult when the subject matter is advanced, especially for second-language learners.

Some specific recommendations for Speech-language Pathologists and teachers of second-language learners may prove helpful. It is important for those working with non-native speakers to consider the mode through which lectures and assignments are delivered. Increasing second-language learners' vocabulary may help with their second language speech perception to the extent that a larger vocabulary may enhance phonetic discrimination skills. In other words, the more words one has in his or her vocabulary, the more necessary it becomes to be able to make the fine phonetic distinctions needed to discriminate among similar sounding words. But the vocabulary instruction ideally

should be done orally rather than through written text and via multiple exemplars of the same word. Relying exclusively on written work deprives the second-language learner the opportunities to hear correct pronunciation of target words and how the word contrasts with similar sounding words. Oral presentation and practice during vocabulary instruction should also involve immediate feedback to heighten the learner's awareness of correct pronunciation and their own mispronunciations.

This study and others like it should be of interest to those who wish to promote intelligibility and comprehensibility in the classroom by incorporating communication strategies that offset the effects of noise when capacity demand is high. Clearly, we need more studies to elucidate the factors that have the greatest impact on intelligibility and comprehensibility for non-native speakers and listeners.

## References

- Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, 55, 64-88.
- Auer, Jr., E. T. & Luce, P. A. (2003). Probabilistic phonotactics in spoken word recognition. In C. T. McLennan, P. A. Luce, G. Maunder, & J. Charles-Luce (Eds.), *University at Buffalo Working Papers on Language and Perception*, Vol. 2, (pp. 164 – 202) Buffalo, NY: University of Buffalo.
- Baddeley, A., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.). *The psychology of learning and motivation: Advances in research and theory*, (pp. 47-89). New York, NY: Academic Press.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, 4, 417-423.
- Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 177-204). Timonium, MD: York Press.
- Bilger, R., Neutzel, J., Rabinowitz, W., and Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, 27, 32-48.

- Boukrina, O., & Marian, V. (2006). Integrated phonological processing in bilinguals: Evidence from spoken word recognition. Proceedings of the Twenty- Eighth Annual Meeting of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum.
- Bracken, B., & McCallum, R. (1998). *Universal Nonverbal Intelligence Test (UNIT)*. Itasca, IL: Riverside Publishing.
- Bradlow, A., & Pisoni, D. (1999). Recognition of spoken words by native and non-native listeners: talker-, listener- and item-related factors. *Journal of the Acoustical Society of America*, 4, 2074-2085.
- Carrow-Woolfolk, C. (1996). *OWLS: Oral and Written Language Scales*. Bloomington, MN: Pearson Assessments.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test-3<sup>rd</sup> Edition Revised*. Circle Pines, MN: American Guidance Service.
- Edwards, J, Fox, R.A., & Rogers, C.L. (2002). Final consonant discrimination in children: Effects of phonological disorder, vocabulary size, and phonetic inventory size. *Journal of Speech, Language, & Hearing Research*, 45, 231-242.
- Ellis, L.W. & Fucci, D. J. (1992) Effects of listeners' experience on two measures of intelligibility. *Perceptual and Motor Skills*, 74, 1099-1104.
- Flege, J. E. & Hillenbrand, J. (1984). Limits on the phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, 76, 708-721.
- Flege, E., & Fletcher, K. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91 (1), 370-389.
- 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). Timonium, MD: York Press.
- Flege, J., Bohn, O. & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25 (4), 437-470.
- Flege, J., Schirru, C. & MacKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40, 467-491.
- Garlock, V., Walley, A., & Metsala, J. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, 45, 468-492.
- Goh, W., & Pisoni, D. (2003). Effects of lexical competition on immediate memory span for spoken words. *The Quarterly Journal of Experimental Psychology*, 56A(6), 929-954.
- Goldinger, S., Luce, P., & Pisoni, D. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, 28, 501-518.
- Goldinger, S., Pisoni, D., & Logan, J. (1991). On the nature of talker variability effects on the recall of spoken word lists. *Journal of Experimental Psychology*, 12(1), 152-162.
- Grosjean, F. (1997). Processing mixed languages: Issues, findings, and models. In A. M. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225-253). Hillsdale, NJ: Erlbaum.
- Imai, Walley, & Flege, J. (2005). Lexical frequency and neighborhood density effects

- on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America*, 117(2), 896-907.
- Kalikow, D., Stevens, K., & Elliot, L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61, 1337-1351.
- Koul, R. K. & Allen, G. D. (1993). Segmental intelligibility and speech interference thresholds of high-quality synthetic speech in presence of noise. *Journal of Speech and Hearing Research*, 36, 790-798.
- Kucera, F., & Francis, W. (1967). Computational analysis of present day American English. Providence, RI: Brown U.P.
- Lewellen, M., Goldinger, S., Pisoni, D., & Greene, B. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General*, 122(3), 316-330.
- Lisker, L., & Abramson, A. (1971). Distinctive features and laryngeal control. *Language*, 47(4), 767-785.
- Lively, S., Logan, J., & Pisoni, D. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94 (3), 1242-1255.
- Lively, S., Pisoni, D., VanSummers, W., & Bernacki, D. (1993). Effects of cognitive workload on speech production: Acoustic analysis and perceptual consequences. *Journal of the Acoustic Society of America*, 93 (5), 2962-2973.

- Logan, J. S. Greene, B. G., & Pisoni, D. B. (1989). Segmental intelligibility of synthetic speech produced by rule. *Journal of the Acoustical Society of America*, 86(2), 566-581.
- Luce, P. (1986). Neighborhoods of words on the mental lexicon. Research on Speech Perception, Technical Report No. 6. Bloomington, IN: Indiana University
- Luce, P., Feustel, T., & Pisoni, D. (1983). Capacity demands on short-term memory for synthetic and natural speech. *Human Factors*, 25, 17-32.
- Luce, P. & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19, 1-36.
- Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language*, 51, 297-306.
- Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology*, 93, 95-119.
- Marian, V., and Blemenfeld, H. (2006). Phonological neighborhood density guides lexical access in native and non-native language production. *Journal of Social and Ecological Boundaries: Special Issue on Bilingualism*, 2, 3-35.
- Mayo, L., Florentine, M., & Buus, S. (1997). Age of second language acquisition and perception of speech in noise. *Journal of Speech and Hearing Research*, 40 (3), 686-693.

- McAllister, R. (1996). Perceptual foreign accent: L2 users' comprehension ability. In A. James & J. Leather (Eds.) *Second language speech structure and process* (pp. 119-132). Berlin: Mouton de Gruyter.
- Meador, D., Flege, J., & Mackay, I. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition*, 3, 55-67.
- Miller, G. & Nicely, P., (1955). Analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338-353.
- Munro, M. & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Available from: <http://www.usf.edu/FreeAssociation/>. Retrieved December, 2006.
- Paris, C., Gilson R., Thomas, M., & Silver, N. (1995). Effect of synthetic voice intelligibility upon speech comprehension. *Human Factors*, 37, 335-340.
- Paris, C., Thomas, M., Gilson, R., & Kincaid, J. (2000). Linguistic cues and memory for synthetic and natural speech. *Human Factors*, 42, 421-431.
- Pastizzo, M. and Carbone, R. (2007). Spoken word frequency counts based on 1.6 million words in American English. *Behavior Research Methods*, 39, 1025-1028.
- Pichura-Fuller, M.C., Schneider, B., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, 97 (1), 593-608.
- Pisoni, D., Nusbaum, H., Luce, P., & Slowiaczek, L. (1985). Speech perception, word recognition, and the structure of the lexicon. *Speech Communication*, 4, 75-95.

- Pisoni, D. & Koen, E. (1981). Some comparisons of intelligibility of synthetic and natural speech at different speech-to-noise ratios. Tech. Rep. 7, Research on Speech Perception Progress Report. Bloomington, IN: Indiana University.
- Pisoni, D., Manous, L., & Dedina, M. (1987). Comprehension of natural and synthetic speech: Effects of predictability on the verification of sentences controlled for intelligibility. *Computer Speech and Language*, 2, 303-320.
- Ralston, J., Pisoni, D., Lively, S., Greene, B., & Mullennix, J. (1991). Comprehension of synthetic speech produced by rule: Word monitoring and sentence-by-sentence listening time. *Human Factors*, 33 (4), 471-491.
- Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Baltimore, MD: York Press.
- Rogers, C., Dalby, T., & Nishi, K. (2004). Effects of noise and proficiency on intelligibility of Chinese-accented English. *Language and Speech*, 47 (2), 139-154.
- Rogers, C., Lister, J., Febo, D., Besing, J., & Abrams, H. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27, 465-485.
- Roodenys, S., Hulme, C., Lethbridge, A., Hinton, M., & Nimmo, M. (2002). Word-frequency and phonological-neighborhood effects on verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(6), 1019-1034.
- Sitler, R. W., Schiavetti, N. & Metz, D. E. (1983). Contextual effects in the measurement

- of hearing-impaired speakers' intelligibility. *Journal of Speech and Hearing Research, 26*, 22-30.
- Sommers, M. (1995). The structural organization of the mental lexicon and its contribution to age-related declines in spoken word recognition. *Psychology and Aging, 11* (2), 333-341.
- Southwood, M.H., & Flege, J.E. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics and Phonetics, 13*, 335-349.
- Stine, E., & Wingfield, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology and Aging, 2*, 272-279.
- Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.) *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-45). Timonium, MD: York Press,
- Strange, W. (1999). Perception of consonants: From variance to invariance. In J. M. Pickett (Ed.), *The acoustics of speech communication: Fundamentals, speech perception theory, and technology* (pp. 166-182). Needham Heights, MA: Allyn & Bacon.
- Strange, W., Bohn, O., Trent, S., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America, 115*(4), 1791-1807.
- Studebaker, G. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, (28)*, 455-462.
- Suenobu, M., Kanzaki, K., & Yamanc, S. (1992). An experimental study of intelligibility

- of Japanese English. *IRAL: International Review of Applied Linguistics in Language Teaching*, 30, 146-156.
- Takata, Y., & Nabelek, A. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America*, 88 (2), 663-666.
- TOEFL iBT Tour (n.d.). Retrieved March 28, 2007 from [http://www.ets.org/Media/Tests/TOEFL/tour/highrez/start-web\\_content.html](http://www.ets.org/Media/Tests/TOEFL/tour/highrez/start-web_content.html)
- Vitevitch, M. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 735-747.
- Walley, A. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, 13, 286-350.
- Werker, J., Fennell, C., Corcoran, K., & Stager, C. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, 3(1), 1-30.
- Wijngaarden, S. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication*, 35, 103-113.
- Wijngaarden, S., Herman, J., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*, 111 (4), 1906-1916.
- Wright, R. (1997). Lexical competition and reduction in speech: A preliminary report. Research on Spoken Language Processing, Progress Report Number 21.  
Bloomington, IN: Indiana University.

## Appendices

Appendix A: Monolingual Language Background Questionnaire

Participant Background Questionnaire

Name: \_\_\_\_\_ Age: \_\_\_\_\_ Address (town & state): \_\_\_\_\_

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: Bilingual Language Background Questionnaire

Participant Background Questionnaire

Name: \_\_\_\_\_ Age: \_\_\_\_\_ Address (town & state): \_\_\_\_\_

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)?  
\_\_\_\_\_

Appendix B: (Continued)

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 readingwriting

Appendix B: (Continued)

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:

b. Comprehension \_\_\_\_\_

c. Fluency (ease of expression) \_\_\_\_\_

d. Vocabulary: \_\_\_\_\_

e. Pronunciation: \_\_\_\_\_

f. Grammar: \_\_\_\_\_

## Appendix C: Easy and Hard Word Lists

### Easy Words

fig live dog  
down move vote  
work food league  
long size thick  
both cause page  
does chief join  
put faith shop  
give pool roof  
young deep leg  
thing firm lose  
peace serve theme  
god reach soil  
five mouth pull  
gave teeth chain  
death gas curve  
shall jack path  
real check dirt  
south king vice  
job shape rough  
love learn balm  
full ship noise  
wife neck thought  
voice watch  
girl judge  
wrong hung

### Hard Words

ban rum pawn  
bead sane bun  
bean soak gut  
bug suck lice  
bum tan mid  
cheer weed hurl  
comb whore moat  
cot wick teat  
den con hash  
dune doom hid  
fade hick hoot  
fin rut mace  
goat toot wad  
knob wade moan  
lad bud mum  
mall dame rim  
mat lace rout  
mitt lame wail  
mole pad hum  
pat chore sill  
pet cod beak  
pup hack hag  
rat kin  
rhyme kit  
chat wed

## Appendix D: Target Words Phonological Neighborhood and Semantic Features Data

Easy Word	Frequency	Neighborhood Density	Neighborhood Frequency	Connectivity	Set Size	Concreteness
balm	36	13	17.77	4.59	17	5.91
both	730	13	22.38			
cause	130	10	33.3	4.86	19	2.9
chain	50	19	39.42	3.36	14	5.85
check	88	15	15.93	4.28	13	4.38
chief	119	12	10.42	2.26	10	4.82
curve	45	13	15.23	4.1	20	4.4
death	277	10	30.7	4.53	15	3.86
deep	109	18	36.17	4.6	13	3.96
dirt	43	15	23.4	7.52	18	5.51
does	485	16	30.5			
dog	75	8	11.875	3.56	5	5.75
down	895	20	38.7	3.78	10	3.23
faith	111	11	50.09	4.29	12	2.71
fig	72	16	44.4375	3.96	10	6.28
firm	109	13	13.69	6.65	20	3.96
five	286	12	46.5	3.34	14	3.53
food	147	11	24.91	5.31	18	5.84
full	230	15	59.53	3.52	11	3.74
gas	98	19	25.68	4.29	10	5.34
gave	285	18	47.67			
girl	220	16	6.69	4.12	8	6.83
give	391	7	70.4286	5.9	13	3.18
god	318	19	77.32	7.07	23	3.61
hung	65	18	30.56	2.85	19	3.88
jack	92	17	74.41		17	5.2
job	238	19	5.32	5.45	7	4.11
join	65	8	27	4.49	20	2.88
judge	77	6	2.33	4.77	17	6.25
king	88	17	36.12	4.7	8	5.54
league	69	19	24.47	4.75	13	
learn	84	16	51	7.46	19	3.66
leg	58	15	79.67	5.06	10	6.04
live	177	15	61.07	6.2	18	4.32
long	755	13	75.85	3.94	13	3.68
lose	58	17	65.76			
love	232	11	42.45	6.45	18	3.51
mouth	103	7	41.86	7.34	19	5.47
move	171	8	16.38	10.19	27	3.81
neck	81	13	15.9231	4.4	19	5.83
noise	37	4	43.5	5.8	14	5.29
page	66	16	52.94	4.07	10	5.85
path	44	14	16.5714	4.17	12	4.93
path	327	19	18.16	3.19	19	2.98
peace	111	18	25.28	3.64	13	6.29
pool	51	16	64.81	4.61	10	3.4
pull	437	14	20	5.41	16	2.77
put	106	20	77.45	3.66	16	3.55
reach	260	16	23.44		14	3.77

Appendix D (continued)

real	59	13	49.69	4.33	9	5.82
roof	41	20	19.5	5.39	17	4.48
rough	107	14	24.79			
serve	267	13	3.85			
shall	85	16	19.31	5.4	18	4.22
shape	83	19	18.2105	5.56	9	6.25
ship	63	16	40.38	5.04	14	5
shop	138	12	71.92	5.41	11	3.5
size	54	13	28.38	2.73	6	5.69
soil	240	5	22.2	2.74	7	3.39
south	103	12	19.33	6.35	16	6.14
teeth	55	8	45.375	5.01	22	3.32
theme	67	13	78.46	2.79	8	3.77
thick	333	11	72		20	3.46
thing	515	11	36.55	3.73	9	1.28
thought	42	16	31.13		13	4.09
vice	226	7	29	5.95	16	5.03
voice	75	15	29.6	6.69	21	3.85
vote	81	5	60.6	6.74	14	4.63
watch	228	15	72.93	3.96	8	5.8
wife	760	20	47	6.12	19	3.88
work	129	13	74.85	3.36	6	2.6

Hard words	Frequency	Neighborhood Density	Neighborhood Frequency	Connectivity	Set Size	Concreteness
ban	1	26	299.19	4.3	20	5.61
bead	1	28	298.21			
beak	5	25	396.36		19	6
bean	9	23	216.48	3.23	10	4.91
bud	4	26	190.58	6.38	16	6.4
bug	7	24	287.38		21	6.08
bum	1	28	410.5357	4.46	12	5.77
bun	5	22	743.36			
chat	8	27	87.22	3.57	21	
cheer	7	21	879.1			
chore	6	27	91.7	3.39	11	6.13
cod	6	24	92.92	4.46	13	
comb	9	21	104.8095		33	5.61
con	1	35	180.37	3.45	11	5.93
cot	7	23	117.52		18	2.38
dame	2	33	130.64			
den	3	23	95.48			
doom	1	27	120.33			
dune	2	22	115.23	5.62	23	2.53
fade	3	30	825.33	5.03	12	
goat	1	24	253.96			
gut	3	30	438.8			
hack	1	25	479.16			
hag	1	22	544.36			
hash	1	25	439.76			
hick	6	25	711.44			

Appendix D (continued)

hid	9	23	146.09	5.09	13	
hoot	5	25	232.52	4.6	11	hoot
hum	3	22	199.18			hum
hurl	2	33	796.27	5.37	12	hurl
kin	2	35	281.86		19	kin
kit	2	21	236.67	2.16	6	kit
knob	7	29	92.34	4.18	24	knob
lace	6	34	187.38			lace
lad	2	28	89.29			lad
lame	2	26	138.31	3.5	9	lame
lice	1	29	175.14			lice
mace	3	24	192.13	4.05	8	mace
mall	8	30	636.03	3.32	12	mall
mat	2	26	113.73			mat
mid	1	33	321.18			mid
mitt	1	26	233.96			mitt
moan	1	31	161.39	4.98	15	moan
moat	4	33	97.39			moat
mole	1	23	144.74			mole
mum	8	26	225.12	5.06	13	mum
pad	35	39	444.72	4.65	19	pad
pawn	2	21	366.95	3.05	11	pawn
pet	8	30	96.63	4.21	11	pet
pup	2	21	98.48			pup
rat	6	37	480.27	4.56	19	rat
rhyme	4	25	121.56	4.14	17	rhyme
rim	5	26	129.31	5.04	20	rim
rout	1	21	164.48			rout
rum	3	29	256.28	6.1	10	rum
rut	1	28	221.21			rut
sane	8	33	90.33			sane
sill	4	35	116.71	2.87	4	sill
soak	7	23	108.91	7.41	16	soak
suck	5	25	142.96		21	suck
tan	9	25	379.28	3.49	13	4.18
teat	1	31	302.74			
toot	3	27	1066.59			
wad	1	22	163.14			
wade	2	24	248.38			
wail	3	32	153.78			
wed	2	25	295.08	3.96	11	3.43
weed	1	24	287	5.02	14	5.96
whore	2	30	689.1			
wick	4	26	432.69	2.26	3	5.45

Appendix E: Practice Words

could  
pond  
ten  
frog  
mop  
dime  
beach  
ran  
gild  
train

## Appendix F: Distracter Words

more

call

take

from

band

grass

## Endnote

<sup>1</sup> Although it is recognized that the frequency norms by Kucera and Francis (1967) are rather old, they are used in the proposed study for several reasons. First, studies which have considered word- frequency use these norms (Bradlow & Pisoni, 1999; Imai et al., 2005; Lewellen et al., 1993; Nelson et al., 1998; Roodenrys et al., 2002; Vitevich, 2002). In order to allow for comparison of results between the proposed study and past studies which have considered word frequency, it was deemed best to use the same norms. Further, I know of no more recent word-frequency norms available.

### About the Author

Astrid Doty received a Bachelor's Degree in Humanities in 1993 and a M.L.A. in Humanities in 1995 from the University of South Florida. She taught as a teaching assistant and later as an adjunct instructor at the University of South Florida in the Department of Humanities. She later went on to earn her M.S. in Communications Sciences and Disorders in 2000 from the University of South Florida and worked as a speech-language pathologist in the public schools for several years before entering the interdisciplinary Ph.D. program in Psychology (Cognitive and Neural Sciences Program) and Communication Sciences and Disorders in 2001.

While in the Ph.D. program at the University of South Florida, Mrs. Doty worked as a graduate teaching assistant and also taught Developmental Psychology. She presented her research at several national and international conferences such as the Acoustical Society of America and the American Speech Language and Hearing Association. Additionally, she worked as a speech-language pathologist in the public school system where she provided therapy for individuals aged three through 21.