WHERE IS THE LOCUS OF DIFFICULTY IN RECOGNIZING FOREIGN-ACCENTED WORDS?
NEIGHBORHOOD DENSITY AND PHONOTACTIC PROBABILITY EFFECTS ON THE RECOGNITION OF FOREIGN-ACCENTED WORDS BY NATIVE ENGLISH LISTENERS

BY

Kit Ying Chan

Submitted to the graduate degree program in Psychology
and the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Chairperson Michael S. Vitevitch, Ph.D.

Susan J. Kemper, Ph.D.

John Colombo, Ph.D.

Joan Sereno, Ph.D.

Allard Jongman, Ph.D.

Date Defended: March 13, 2012

The Dissertation Committee for Kit Ying Chan
certifies that this is the approved version of the following dissertation:

WHERE IS THE LOCUS OF DIFFICULTY IN RECOGNIZING FOREIGN-ACCENTED WORDS?
NEIGHBORHOOD DENSITY AND PHONOTACTIC PROBABILITY EFFECTS ON THE RECOGNITION OF FOREIGN-ACCENTED WORDS BY NATIVE ENGLISH LISTENERS

Chairperson Michael S. Vitevitch, Ph.D.

Date approved: April 17, 2012

Abstract

This series of experiments (1) examined whether native listeners experience recognition difficulty in all kinds of foreign-accented words or only in a subset of words with certain *lexical* and *sub-lexical* characteristics— neighborhood density and phonotactic probability; (2) identified the locus of foreign-accented word recognition difficulty, and (3) investigated how accent-induced mismatches impact the lexical retrieval process. Experiments 1 and 4 examined the recognition of native-produced and foreign-accented words varying in neighborhood density with auditory lexical decision and perceptual identification tasks respectively, which emphasize the lexical level of processing. Findings from Experiment 1 revealed increased accent-induced processing cost in reaction times, especially for words with many similar sounding words, implying that native listeners increase their reliance on top-down lexical knowledge during foreign-accented word recognition. Analysis of perception errors from Experiment 4 found the misperceptions in the foreign-accented condition to be more similar to the target words than those in the native-produced condition. This suggests that accent-induced mismatches tend to activate similar sounding words as alternative word candidates, which possibly pose increased lexical competition for the target word and result in greater processing costs for foreign-accented word recognition at the lexical level. Experiments 2 and 3 examined the sub-lexical processing of the foreign-accented words varying in neighborhood density and phonotactic probability respectively with a same-different matching task, which emphasizes the sub-lexical level of processing. Findings from both experiments revealed no extra processing costs , in either reaction times or accuracy rates, for the foreign-accented stimuli, implying that the sub-lexical processing of the foreign-accented words is as good as that of the native-produced words. Taken together, the overall recognition difficulty of foreign-accented stimuli, as well as the differentially increased processing difficulty for accented dense words (observed in Experiment 1), mainly stems from the lexical level, due to the increased lexical competition posed by the similar sounding word candidates.

Acknowledgments

I owe my deepest and most sincere gratitude to my advisor, Dr. Michael Vitevitch for his wide knowledge, patience, personal guidance, and unlimited support during my graduate career. I would also like to thank my dissertation committee members, Dr. Susan Kemper, Dr. John Colombo, Dr. Joan Sereno, and Dr. Allard Jongman, for their valuable input and guidance for my dissertation. I wish to express my sincere thanks to Professor Hsuan Chih Chen, who introduced me to the field of Cognitive Psychology and guided me through my first steps into Psycholinguistic studies. I owe my loving thanks to my family and friends for their unending encouragement, support, and faith in me.

Finally, the financial support of the National Institutes of Health-National Institute of Deafness and Other Communication Disorders (NIH-NIDCD R01 DC 006472) is gratefully acknowledged.

Table of Contents

List of Figures

List of Tables

List of Appendix

Introduction

In order to be competitive in today's globalized world, the capability to speak more than one language becomes a necessary skill to have, and there are a growing number of multilingual speakers. Using English as an example, as of the year 2000, 375 million people speak English as a second language, and around 750 million people speak English as a foreign language for communication with people from other countries (Graddol, 1997). In daily communication situations, it is unavoidable to encounter interlocutors speaking in a language other than their native language with foreign accents.

In this article, the term *accent* refers to foreign accent— the extent to which the pronunciation of a second language learner is perceived to differ from the native speaker norms (Munro & Derwing, 1995a). Foreign accents are often reported to induce a variety of comprehension difficulties in native listeners, resulting in communication costs for both the speakers and listeners (Gill, 1994; Munro & Derwing, 1995a, 1995b; Schmid & Yeni-Komshian, 1999). For instance, compared to native speech, foreign-accented speech is generally less intelligible (Munro & Derwing, 1995a), requires a longer processing time (Munro & Derwing, 1995b), and is more vulnerable to adverse effects of noise on its intelligibility (Lane, 1963; Munro & Derwing, 1998; Van Wijngaarden, 2001). Also, mispronunciations in accented speech are detected less accurately and more slowly (Schmid & Yeni-Komshian, 1999).

The language barrier created by foreign accents could interfere with communication and pose different consequences in different settings. In an educational context, teachers' foreign accents might hinder students' comprehension of lecture information and have a negative effect on their learning process and educational experience (Gill, 1994). In a business context, employees' foreign accents could lead to customer frustration and economic loss, especially when it comes to customer service and technical support outsourced overseas. For example, customer dissatisfaction with foreign accents is such a big concern that companies pay a premium wage for foreign workers with dialect-neutralized speech (Stafford, 2009).

The topic of foreign-accented speech has started to attract more attention from psycholinguistic researchers in recent years. Numerous studies have addressed some of the important preliminary problems, such as the properties of accented speech and their impacts on speech recognition and judgment of degree of accentedness, as well as how perceptual learning enables listeners to acclimate to accented speech. However, these studies have exclusively focused on how the different acoustic-phonological deviations induced by foreign accents impair foreign-accented speech perception as a whole. Currently, little is known about precisely how foreign accents impact the different stages of spoken word recognition and lead to comprehension costs in native listeners.

Many important questions still remain to be answered: How does the native listeners' spoken word recognition system deal with mismatches between accented speech input and their native phonological representations? Where is the locus of the processing difficulty? Does recognition of foreign-accented words rely more on bottom-up speech signals or top-down lexical knowledge? Is word recognition differentially influenced by sub-lexical and lexical factors, such as phonotactic probability and neighborhood density? Since these questions have yet to be addressed, how foreign accents impair spoken word recognition and how current models of spoken word recognition account for it are still the signature problems in the field.

The overall goal of the current research is to examine the impact of foreign accents on the sub-lexical and lexical stages of word recognition in native listeners. The specific aims of the current experiments are (a) to investigate whether native listeners experience recognition difficulty in all kinds of foreign-accented words or only in a subset of words with certain *lexical* and *sub-lexical* characteristics—neighborhood density and phonotactic probability; (b) to locate the locus of accented word recognition difficulty, and (c) to investigate how accent-induced mismatches impact the lexical retrieval process specifically.

Through comparing the sub-lexical and lexical processing of native and foreign-accented speech, results from the current experiments could also give us insights into the level of processing through which foreign accents induce recognition difficulties. The findings will deepen our understanding of the nature of the processing difficulties experienced by native listeners. Before discussing the testable hypotheses and predictions in detail, the properties of foreign-accented speech and the problems it poses to the native listeners' spoken word recognition system are briefly introduced below, followed by a literature review of research relating to foreign-accented word recognition.

Properties of Foreign-accented Speech

Compared to native speech, foreign-accented speech is characterized by a combination of deviances in subsegmental, segmental, and suprasegmental levels (Flege, 1984). At the subsegmental level, deviances were observed in voice onset time (VOT) difference in stop consonants (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973), and formant frequencies and vowel durations in vowels (Munro, 1993). For instance, Spanish-accented English /t/ has significantly shorter VOT values than that produced by native speakers of English (Flege, 1991).

At the segmental level, mispronunciations always involve substitution or distortion of consonants or vowels (Flege & Hillenbrand, 1984; Munro & Derwing, 1995a). For example, the /r/ in “rice” is substituted by a /l/ in foreign-accented English. Syllable structure errors are also common, which often involve the addition or deletion of a segment or syllable, or the reordering of segments in syllables. Complex syllables, such as CVC, are often reduced to simpler CV syllables by L2 speakers through consonant deletion and vowel insertion (Tarone, 1980). At the suprasegmental level, L2 learners also show deviations in prosody, including stress and intonation patterns, as well as phrasing and rhythm (Anderson-Hsieh, Johnson, & Koehler, 1992; Reed, 2000; Riazantseva, 2001; Temple, 2000). For example, incorrect syllables are stressed in words; intonation contours are inappropriate; pauses do not occur at syntactic boundaries; and stressed syllables are not sufficiently prominent (Anderson-Hsieh et al., 1992).

On top of the absolute deviation from native-like pronunciations, there is a greater acoustic variability within-speakers and across-speakers with foreign accents (Van Compernolle, 2001). This was evidenced in a study by Wade et al. (2007), which compared the vowels in a set of monosyllable words produced by Spanish-accented and native English speakers respectively. In addition to the difference observed in absolute location, all of the Spanish-accented vowels examined (except /æ/) were significantly more variable in observed height and backness values than the native vowels.

Why is Foreign-accented Speech so Challenging for Native Listeners?

Speech perception involves the mapping process linking heard speech signals to lexical representations so that spoken words can be identified and the meaning of the entire utterances can be derived. Current spoken word recognition models generally assume that the speech signal is transformed into prelexical representations, such as features, phonemes, and syllables, prior to lexical access. The accent-related acoustic-phonetic deviations induce a mismatch between accented speech inputs and the listeners' native phonological representations, which is crucial in determining the success of prelexical matching and thus subsequent lexical access. Numerous studies have shown that acoustic-phonetic deviations in foreign-accented speech influence native listeners' perceptions of foreign accents and their understanding of the messages delivered (Anderson-Hsieh et al., 1992; Derwing & Munro, 1997; Koster & Koet, 1993; Magen, 1998; Munro & Derwing, 1995a; Schairer, 1992; Tajima, Port, & Dalby, 1997).

Moreover, due to this greater acoustic variability in foreign-accented speech, the nonnative sound categories tend to be distributed in much greater proximity than the native categories, and this increases category overlap and leads to confusability (Sidaras, Alexander, & Nygaard, 2009; Wade et al., 2007). For example, in the above-mentioned study by Wade and colleagues (2007), certain neighboring Spanish-accented vowel pairs, such as the /i/ and /ɪ/ pair, and the /u/ and /ʊ/ pair, were found to locate closer together in vowel space than their native counterparts with more overlapping. When discriminant analysis was used to predict the confusability of these vowel categories, all Spanish-accented categories were on average about 10% more confusable than the native categories, except the /æ/ and /ɛ/ categories (Wade et al., 2007). The two most confusable vowels /i/ and /ʌ/ according to the discriminant analysis were confirmed to pose the most recognition errors when the native English listeners were trained to identify isolated Spanish-accented English words in another experiment in that study (Wade et al., 2007).

Similar findings were noted in a study by Sabrina and colleagues (2009), in which native English speakers were trained with English words and sentences produced by Spanish-accented learners of English. More temporal (vowel duration) or spectral (F1 and F2 format center frequency) overlap was found between the Spanish-accented English vowels, such as /i/ and /ɪ/, which were more confusable to the native listeners. These results suggest that foreign accents disrupt prelexical processing, which subsequently contributes to the word recognition difficulties in native listeners. Is disrupted pre-lexical processing the major factor contributing to accented word recognition difficulty? How would it affect later processing?

Where does the Locus of Processing Difficulty Lie?

Many of the current models of spoken word recognition (Lahiri & Marslen-Wilson, 1991; McClelland & Elman, 1986; Norris, 1994) adopt an abstractionist view of lexical access. That is, lexical entries consist of a set of abstract, ideal, and modality-free representations. Therefore, the perception system is assumed to filter and discard the surface details tangential to the word identity through normalization, leaving only canonical mental representations at the prelexical level for subsequent lexical processing. For foreign-accented speech in particular, an accent normalization mechanism is assumed to remove all the pronunciation deviations due to foreign accents before lexical processing. Thus, foreign accents should probably only affect sub-lexical processing directly such that prelexical perception is calibrated according to the phonemic categories of the foreign-accented speaker before lexical matching.

As the traces of foreign accent are assumed to be filtered out during accent normalization and should not enter the stage of lexical processing, it is predicted that inadequate prelexical processing resulting from foreign accents does not affect lexical access directly. However, it is unquestionable that lexical processing is affected indirectly by the decisions made at the sub-lexical levels. After prelexical processing, lexical access is generally implemented in most current models of spoken word recognition with some form of explicit or implicit activation and competition among the word candidates in the mental lexicon. As the speech input unfolds so that the amount of evidence from the acoustic-phonetic input accumulates, multiple lexical candidates are activated and compete for recognition.

When any phonological mismatch occurs with the incoming speech, the activated potential lexical candidates fade back into their resting states. This could happen more often especially for foreign-accented speech, due to the greater mismatch induced by the acoustic-phonetic deviations at the prelexical level. That means, the target word may not be as activated as in native speech; therefore longer time may be needed for activation to pass the threshold for recognition in the presence of foreign accents. Also, more lexical candidates might be activated for foreign-accented inputs given that the nonnative sound categories are acoustically more variable and distributed in much greater proximity. With more word candidates competing with the target words, it might take a longer time for activation and competition to be resolved to retrieve the best match for the accented input. It might also be more likely for one of the competing candidates to be mistakenly recognized as the target words. Therefore, the difficulty of foreign-accented word recognition may not be exclusively located at the pre-lexical level; processing at the lexical level may also be less efficient.

The accent-related processing costs have been extensively investigated through a variety of paradigms in numerous studies and manifested differently in terms of lower intelligibility, longer processing time, and higher vulnerability to noise (Clarke

& Garrett, 2004; Lane, 1963; Munro, 1998; Munro & Derwing, 1995a, 1995b; Schmid & Yeni-Komshian, 1999; Van Wijngaarden, 2001). However, these studies mostly use experimental tasks that require word recognition, or emphasize listeners' understanding of the messages delivered in the speech, such as a transcription task, a sentence-verification task, and a listening for mispronunciation task. These experimental tasks emphasize lexical level of processing, but sub-lexical level of processing is generally required before lexical access. Thus, the experimental tasks used in these studies could not allow us to infer whether the processing costs actually originate from the sub-lexical level, the lexical level, or both.

Moreover, these studies have almost exclusively utilized sentence stimuli or stimulus words embedded in carrier sentences so that the influences from the higher-level semantic/syntactic knowledge and the lower-level acoustic-phonetic deviations were not separable. Also, the processing time for individual words could not be deduced from the reaction times measured for the sentence-length stimuli. Even if that is made possible by using isolated words as stimuli, the reaction times only reflect the progressive processing time of both the sublexical and lexical levels. Taken together, most previous research has only demonstrated that foreign accents induce processing costs on word recognition in general, but it has not shed any light on how foreign accents impact the different stages of processing during spoken word recognition. Hence, the goal of the current study is to use isolated word and nonword stimuli with time-sensitive tasks that emphasize sub-lexical and lexical processing respectively to isolate the effects of foreign accent on the sub-lexical and lexical levels of processing.

Before further discussing the experiments in the current study, previous research studying lexical and sub-lexical factors on foreign-accented speech processing are reviewed first.

Influences of Lexical Information on Foreign-accented Speech Processing

Little research has investigated the possible contribution of inefficient sub-lexical and lexical processing to the increased processing costs for foreign-accented word recognition in native listeners. However, there is evidence showing that native listeners experienced increased difficulty in sub-lexical processing of foreign-accented speech and used top-down postlexical information for compensation, at least under adverse listening conditions (Burki-Cohen, Miller, & Eimas, 2001).

In a series of experiments, Burki-Cohen and colleagues (2001) adopted a task manipulation that promoted sub-lexical processing in one condition and lexical processing in another condition to investigate how native listeners use prelexical and lexical information differently in processing of moderately Swiss German-accented and native English words. A phoneme monitoring procedure that requires only sub-lexical processing for optimal performance, was used as a function of a secondary task that promotes lexical processing (this paradigm was first used by Eimas, Hornstein, & Payton, 1990). In that phoneme monitoring task, native listeners were instructed to indicate as quickly and as accurately as possible whether the monosyllablic word presented began with a target phoneme. In some of the experiments, a secondary task was included to prompt lexical processing, in which listeners were asked to respond as quickly as possible whether the target-bearing word was a noun or verb. Half of the stimulus words had high word frequency and the other half had low word frequency. There were three factors: accent type (native vs. accented), secondary task (present vs. absent), and word frequency (high vs. low).

Results showed an overall increasing trend in reaction time for the accented words compared to the native words, but no reliable main effect of accent type or any interactions with accent type was found under ideal listening conditions. In the absence of the secondary task, no reliable frequency effects were found in the phoneme monitoring reaction times, regardless of whether the words were native-produced or foreign-accented. When the phoneme monitoring task included the secondary task, a reliable frequency effect was found for both native-produced and foreign-accented stimulus words. This suggested that phonemic decisions for both native-produced and foreign-accented words primarily relied on prelexical information in the absence of a secondary task. These same decisions primarily relied on postlexical information when a secondary task, which emphasizes lexical access, was included. That is, the phoneme monitoring reaction time measures only pre-lexical processing without the secondary task, but it also measures lexical processing when followed by the secondary task. The failure to detect any processing costs for the sub-lexical and lexical processing of foreign-accented words could be due to the use of a relatively moderate German accent that was clearly perceptible to native listeners in this study.

When the listening conditions were degraded by adding multitalker babbling noise, contrasting patterns emerged. First, the phoneme monitoring reaction times were significantly longer for the foreign-accented than for the native-produced words, showing that there is an increased difficulty in sub-lexical processing of foreign-accented words. More importantly, the overall pattern of results was maintained for native-produced words in noisy listening conditions: a significant frequency effect on phoneme monitoring reaction times was found only in the presence of the secondary task; however, for foreign-accented words embedded in noise, a significant frequency effect on phoneme monitoring reaction times was found even in the absence of the secondary task. This result suggested that under degraded listening conditions, listeners experience increased difficulty in sub-lexical processing of foreign-accented words which they compensate for by using top-down lexical knowledge, as opposed to prelexical information, for phonemic processing of foreign-accented words.

Overall, the results from this series of experiments showed that the processing of foreign-accented speech appeared to be different from that of native speech regarding the use of prelexical versus top-down postlexical information, at least under degraded listening conditions. If some sort of top-down lexical information is required to compensate for the disrupted pre-lexical processing during foreign-accented word recognition, lexical effects are expected to influence the recognition of foreign-accented words. For instance, lexical effects, including lexical frequency and neighborhood density, which have been shown to affect recognition of native speech (Luce & Pisoni, 1998; Vitevitch, 2002; Vitevitch & Rodriguez, 2004), should be considered. Hence, in the following section, we will review previous research examining the influence of a lexical factor, neighborhood density, on foreign-accented word recognition.

Neighborhood Density Effect on Accented Word Recognition

To date, few studies have explored the influence of lexical factors on foreign-accented word recognition (Imai, Walley, & Flege, 2005; Levi, Winters, & Pisoni, 2007). In most of the previous studies on foreign-accented speech recognition, the lexical characteristics of the stimulus words, such as word frequency and neighborhood density, have not been controlled or systematically manipulated (e.g., Clarke & Garrett, 2004; Lane, 1963; Munro & Derwing, 1995a, 1995b; Schmid & Yeni-Komshian, 1999; Van Wijngaarden, 2001). As shown in Burki-Cohen and colleagues (2001), if native listeners do tend to rely on top-down lexical knowledge to compensate for the disrupted pre-lexical processing during foreign-accented word recognition, lexical factors are expected to influence the recognition of foreign-accented words.

Consider the lexical factor— neighborhood density, which refers to the number of words that are phonologically similar (i.e., phonological neighbors) to a target word. Words with many similar sounding neighbors are said to have dense neighborhoods, whereas words with few similar sounding neighbors are said to have sparse neighborhoods. Based on the Neighborhood Activation Model (NAM) of spoken word recognition, the recognition of a spoken word depends on its phonological similarity to other words in the mental lexicon (Luce & Pisoni, 1998). Due to a large number of confusable competitors, the recognition of words from dense neighborhoods relies more on fine phonetic discriminations at the segmental level than the recognition of words from sparse neighborhoods. Therefore, during foreign-accented word recognition, it is expected that words from dense neighborhoods, which require more fine discrimination at the prelexical level, would be harder to recognize than words from sparse neighborhoods, due to the less efficient sub-lexical processing for accented speech.

Imai and colleagues (2005) examined the influence of lexical frequency and neighborhood density on the recognition of native-produced and Spanish-accented words by native English and native Spanish listeners. Listeners were asked to identify words embedded in multitalker babbling noise. Many other variables were also studied in this experiment, but only results related to our current discussion— native listeners' recognition of foreign-accented words varying in neighborhood density— will be discussed here. A significant neighborhood density effect was found only in the recognition of foreign-accented words. Spanish-accented sparse words were recognized more accurately than Spanish-accented dense words, whereas no such difference was observed for the native-produced words.

This finding supports the idea that foreign accents have greater impairments on the recognition of words from dense neighborhoods rather than words from sparse neighborhoods, as a fine-grained phonological discrimination is required for distinguishing the dense target words from the competing similar sounding words. The neighborhood density effect observed in the foreign-accented condition suggests that the native listeners relied heavily on lexical information for foreign-accented word recognition. Consistent with Burki-Cohen and colleagues' (2001) results, native listeners tend to use top-down lexical information to compensate for the sub-optimal inputs from foreign-accented speech.

It is worthy to note the limitations in this study by Imai and colleagues (2005). First, recognition accuracies from this intelligibility task only measure the end-product of the entire process of spoken word recognition; it only allows us to induce that there are processing costs somewhere along the spoken word recognition process. As with most of the previous research, this study could not determine the locus of the processing difficulty for foreign-accented word recognition. Moreover, this study used noise-degraded stimuli. Given that noise has been shown to have more adverse effects on foreign-accented speech (Munro, 1998), it is possible that the lexical effects are amplified in the presence of noise and might not be found in ideal listening conditions. It is important to examine whether the lexical effects on foreign-accented word recognition could be extended to *noise-free* stimuli. Thus, to overcome these limitations, Experiment 1 in the current study used a time-sensitive task with noise-free stimuli to examine whether the enhanced neighborhood density effect on foreign-accented word recognition is extended to noise-free stimuli and manifested in reaction times.

Sub-lexical Factors on Accented Word Recognition

Apart from top-down lexical information, sublexical information that constrains the sequences and the segmental co-occurence relations in syllables may also be used during foreign-accented word recognition. Phonotactic probability, which refers to the relative frequencies of positional segments and biphones, has been demonstrated to influence native-spoken word recognition in previous research (Vitevitch, 2003; Vitevitch & Luce, 1998, 1999).

In Vitevitch's study (2003), phonotactic probability was measured by (1) how often a particular segment occurs in a given position in a word (positional segment frequency), and (2) how often two particular segments co-occur in sequence in a word (biphone frequency). The set of stimulus words used varied in phonotactic probability and neighborhood density. Due to the positive correlation between neighborhood density and phonotactic probability, words comprised of common segments and sequences of segments tend to have many similar sounding neighbors, whereas words comprised of less common segments and sequences of segments tend to have few similar sounding neighbors.

A same-different matching (a.k.a. AX) task was used, in which participants were presented with two spoken stimuli in a row and required to respond as quickly and as accurately as possible if the two items were the same or different. In contexts with different proportion of words and nonwords as filler items, the same set of stimulus words varying in phonotactic probability/ neighborhood density showed opposite patterns of results. When mostly nonsense words were used as filler items to encourage the use of sub-lexical representations for processing, a phonotactic probability effect was observed—high probability/dense words were responded to more quickly than the low probability/sparse words. When mostly real words were used as filler items to encourage the use of lexical representations for processing, a neighborhood density effect was observed for the same set of stimulus words. That is, high probability/dense words were responded to more slowly than low probability/sparse words.

These results suggest that although listeners are dominantly influenced by neighborhood density during lexical processing, they are sensitive to phonotactic information during sub-lexical processing of native-produced spoken words. Although, the lexical level of processing is typically more dominant than the sublexical level during real word recognition (Vitevitch & Luce, 1999), information regarding the probability of phonotactic patterns does influence processing of spoken words for native-produced speech.

It is possible that phonotactic probability may play a more important role in foreign-accented word recognition. Given that accented speech produces severe mismatches at the sub-lexical level, native listeners might have to rely more heavily on sub-lexical representations to retain that sequence of sounds until a matching lexical representation can be retrieved. Thus, native listeners may rely more heavily on phonotactic information for restoring foreign-accented speech, resulting in a phonotactic probability effect on foreign-accented word recognition. However, to the best of my knowledge, there is no previous research on the effect of phonotactic probability on foreign-accented word recognition. Therefore, this topic was examined in the current study, which is described in more detail below.

Overview of the Current Study

Previous studies have shown that native listeners have difficulty recognizing foreign-accented words (Gill, 1994; Munro & Derwing, 1995a, 1995b; Schmid & Yeni-Komshian, 1999). This difficulty is probably due to the acoustic-phonetic mismatches induced by foreign accents. To date however, little is known about precisely how foreign accents impact the different stages of spoken word recognition processing, and thus lead to the processing costs in native listeners. Most previous studies utilized sentence stimuli or stimulus words embedded in carrier sentences to study the processing costs in accented word recognition (Clarke & Garrett, 2004; Lane, 1963; Munro, 1998; Munro & Derwing, 1995a, 1995b; Schmid & Yeni-Komshian, 1999; Van Wijngaarden, 2001). The influence of higher-level semantic/syntactic knowledge makes it difficult to assess the impact of foreign accents on spoken word recognition processes in these studies.

To overcome the limitations of previous studies, the use of isolated word stimuli with time-sensitive tasks that emphasize the sub-lexical and lexical processing respectively was used in the current study to assess the time-course of the effects of foreign accents during processing. The findings of Imai and colleagues (2005) regarding the influence of neighborhood density on foreign-accented word recognition, as well as the findings of Vitevitch (2003) regarding the influence of phonotactic probability on sub-lexical processing of native-spoken words, provides the impetus for choosing these two variables as the lexical and sub-lexical factors to focus on in the current study.

The overall goal of the current study is to examine the impact of foreign accents on the sub-lexical and lexical stages of word recognition in native listeners. The specific aims of the experiments were (a) to investigate whether native listeners experience recognition difficulty in all kinds of foreign-accented words or only in a subset of words with certain *lexical* and *sub-lexical* characteristics— neighborhood density (Experiment 1) and phonotactic probability (Experiment 3); (b) to identify the locus of accented word recognition difficulty (Experiment 2), and (c) to investigate how accent-induced mismatches impact the lexical retrieval process (Experiment 4).

Like Imai and colleagues (2005), Experiment 1 examined whether foreign-accented word recognition is undermined for all words or only for a subset of words with certain *lexical* characteristics— neighborhood density. Unlike Imai's et al study (2005), this experiment used *noise-free* stimuli to check whether the effect of neighborhood density on foreign-accented word recognition extends to an ideal listening condition. Most importantly, a time-sensitive task— lexical decision task— was used to assess the processing costs of foreign-accented word recognition in terms of reaction times, in addition to accuracy rates.

Spoken word recognition involves processing at both the sub-lexical and lexical level. Experiment 2 aimed to examine whether the accent-induced processing costs is also observed at the sub-lexical level. With the same set of dense and sparse stimulus words from Experiment 1, a same-different matching task that is time-sensitive and emphasizes processing at the sub-lexical level was used in Experiment 2. Results from Experiment 2 could demonstrate whether native listeners also experience difficulty in sub-lexical processing of foreign-accented words in general and foreign-accented words from dense neighborhoods in particular. More importantly, by comparing the processing costs, as manifested as increased reaction times compared to native speech, across tasks from Experiments 1 and 2, we could distinguish the processing costs originating from the sub-lexical and lexical levels respectively. Specifically, it gives us insights into whether the locus of increased processing difficulty in foreign-accented dense words lies on the sub-lexical or lexical level.

In addition to lexical factors, Experiment 3 examined whether foreign-accented word recognition is also influenced by the sub-lexical characteristic— phonotactic probability. A speeded same-different matching task was used in Experiment 3 to investigate the sub-lexical processing of a set of foreign-accented words and nonwords varying in phonotactic probability. Processing costs were assessed in terms of both reaction times and accuracies rates. Results from Experiment 3 could indicate whether native listeners rely more heavily on phonotactic information in the sub-lexical processing of foreign-accented stimuli compared to native-produced stimuli. It will also give us insight into whether native listeners make use of sub-lexical information to compensate for the distorted foreign-accented speech inputs.

To more closely examine the impact of foreign accents at the lexical retrieval process, Experiment 4 used an intelligibility task to collect the perception errors (misperceptions) resulting from foreign accents, as well as from noise in native speech. The misperceptions collected in this task were analyzed and compared with the target words in terms of phonological similarity. The misperception analysis could reveal the set of lexical candidates being activated during the lexical competition. Thus, the results shed light on how specifically the mismatches induced by foreign accents lead to increased processing difficulty at the lexical level.

The effect of perceptual adaption to foreign-accented speech was also examined in Experiments 1 in the current study. Previous research showed that with a brief exposure to utterances produced by multiple talkers with the same foreign accent, native listeners perceptually adapted to accent-general systematic variations, which facilitated the subsequent recognition of non-native speech produced by novel speakers with the same foreign accent (Bradlow & Bent, 2008; Sidaras et al., 2009). Although the benefit effect of accent-level learning found tended to be small and was not unchallenged (cf. Wade et al., 2007), it might be affecting the spoken word recognition performance over the course of the whole experiment. If the listeners are really adapting to the foreign accents in Experiment 1, it would be interesting to see whether perceptual learning interacts with the lexical factor— neighborhood density. That is, would the benefit effect of perceptual learning differ for foreign-accented dense and sparse words?

The participants' experience in learning Spanish and listening to Spanish-accented speech was also reported in the *listeners* section even though different studies reported mixed results regarding the influence of native listeners' experience with foreign-accented speech. For example, Munro, Derwing, and Morton (2006) revealed no advantage for the native listeners in understanding speech spoken with foreign accents that they are more familiar with. However, some other studies have shown that native listeners with extensive exposure to L2 speech are more accurate than listeners with little exposure to L2 speech at transcribing sentences spoken with foreign accents (Bradlow & Bent, 2008; Kennedy & Trofimovich, 2008).

Experiment 1

Native listeners tend to rely on top-down lexical knowledge to compensate for the disrupted pre-lexical processing during foreign-accented word recognition (Burki-Cohen et al., 2001). In line with this idea, not all the foreign-accented words are difficult for native listeners to recognize. Instead, words with certain lexical characteristic are more difficult to recognize in the presence of foreign accents. Imai and colleagues (2005) found a significant neighborhood density effect in the recognition of foreign-accented words, but not in the recognition of native-produced words. In that study, Spanish-accented sparse words were recognized more accurately than Spanish-accented dense words, whereas no such difference was observed for the native-produced words. Foreign-accented dense words have greater processing costs, because a fine-grained phonological discrimination is required for distinguishing the dense target words from the competing similar sounding words.

This result was based on recognition accuracies from an intelligibility task using noise-degraded stimuli. Recognition accuracy is the end-product of the spoken word recognition process; it does not provide any way for us to induce whether the processing cost actually stems from the sub-lexical or lexical level of processing. Moreover, the presence of noise has been shown to result in a larger decrement in intelligibility for foreign-accented speech than for native speech (Munro, 1998). The neighborhood density effect might be amplified in the presence of noise in Imai et al.'s study (2005) and might not be found in ideal listening conditions. Thus, Experiment 1 aimed to overcome these limitations by using a time-sensitive task and noise-free stimuli to further examine whether the neighborhood density effect on foreign-accented word recognition could be extended to ideal listening conditions and replicated in a different paradigm.

An auditory lexical decision task was used in Experiment 1 to assess the time course of the neighborhood density effect on the recognition of foreign-accented isolated words. In the task, participants were presented with either a word or a nonword without any noise over a set of headphones. Participants were asked to decide as quickly and as accurately as possible whether the given stimulus is a real word in English or a nonsense word. Reaction times and accuracy rates were measured as dependent variables. Previous studies, which demonstrated the increased processing time for foreign-accented speech relative to native speech (Clarke & Garrett, 2004; Munro & Derwing, 1995b; Schmid & Yeni-Komshian, 1999), only used sentence-length stimuli. Without a sentence context, the participants in the present experiment cannot use any semantic/syntactic cues for word recognition. Therefore, the genuine impacts of foreign accents on word recognition can be examined in the present experiment.

If foreign accents affect the lexical level of processing during spoken word recognition, foreign-accented words varying in neighborhood density are expected to show different processing costs. Predicting based on findings from Imai et al.'s (2005) study, it is hypothesized that foreign-accented dense words would show a greater processing cost than sparse words. Furthermore, if the mismatches induced by foreign accents drive the native listeners to rely more heavily on the lexical information to resolve the ambiguity in the accented speech signals, the neighborhood density effect should be further enhanced in the presence of foreign accents. That is, words from dense neighborhoods should be responded to more slowly than words from sparse neighborhoods, especially in the foreign-accented condition. Previous studies in native speech normally found no difference in accuracy rates in a lexical decision task even when a significance difference was found in reaction times (Vitevitch & Luce, 1999). Thus, accuracy rates were not expected to be different for the dense and sparse accented words in Experiment 1.

To examine the possible influence of perceptual adaptation to foreign-accent on words varying in neighborhood density in this experiment, the presentation of the stimuli was divided into two blocks and performance in the two blocks were checked for any effect of perceptual learning. A general perceptual adaption to foreign accents was expected such that listeners perform better in the second block relative to the first block. However, no clear prediction can be made whether there would be a differential perceptual learning effect for dense and sparse words.

Method

Stimuli and design

The 64 English monosyllabic stimulus words used in the present experiment all contained three phonemes, in a consonant-vowel-consonant structure. Half of the stimuli had a dense neighborhood density, and half had a sparse neighborhood density. These two groups of stimulus words and their lexical characteristics are listed in Appendix A.1 and A.2 and further described in the following sections.

Neighborhood Density. Neighborhood density measures the number of words that are phonologically similar (i.e., phonological neighbors) to the target words. A word is considered a phonological neighbor of the target word if it differs from the target word by one phoneme substituted, deleted, or added into any position (Greenberg & Jenkins, 1967; Landauer & Streeter, 1973; Luce & Pisoni, 1998). For example, the word *cat* has as phonological neighbors: *_at, scat, rat, cut* and *cap*. Note that *cat* has other neighbors, but only a few were listed for illustration. The neighborhood density value for each stimulus was obtained from a Web-based calculator described in Storkel and Hoover (2010). A group of dense words and sparse words was each selected for the present study under the constraint that subjective familiarity, word frequency, neighborhood frequency, phonotactic probability, and distribution of phonemes (further described later) were equivalent between the two conditions. The selected dense words had a mean neighborhood density value of 27.44 (*SEM* = .351), and sparse words had a mean neighborhood density value of 15.97 (*SEM* = .466). The difference between the two groups of stimuli was significant, *F* (1, 62) = 387.36, *p* < .0001.

Subjective familiarity. Subjective familiarity was measured on a seven-point scale (Nusbaum, Pisoni, & Davis, 1984). Words from dense neighborhoods had a mean familiarity value of 6.93 (*SEM* = .031), and words from sparse neighborhoods had a mean familiarity value of 6.87 (*SEM* = .044, *F* (1, 62) = 1.23, *p* > .05), indicating that all of the words were highly familiar.

Word frequency. Word frequency refers to the average occurrence of a word in the language. Average log word frequency (log-base 10 of the raw values from Kučera & Francis, 1967) was 1.27 (*SEM* = .131) for the dense words and 1.30 (*SEM* = .123) for the sparse words, *F* (1, 62) < 1.

Neighborhood frequency. Neighborhood frequency is defined as the mean word frequency of the neighbors of the target word. Words from dense neighborhoods had a mean log neighborhood frequency value of 3.59 (*SEM* = 1.570) and words from sparse neighborhoods had a mean log neighborhood frequency value of 2.02 (*SEM* = .043, *F* (1, 62) = 1.008, *p* > .05).

Phonotactic probability: The phonotactic probability is measured by how often a certain segment occurs in a certain position in a word (positional segment frequency) and the segment-to-segment co-occurrence probability (biphone frequency; Vitevitch & Luce, 1998). The mean positional segment frequency for dense and sparse words were .152 (*SEM* = .005) and .146 (*SEM* = .007) respectively, *F* (1, 62) < 1. The mean biphone frequency for dense and sparse words were .007 (*SEM* = .0007) and .006 (*SEM* = .0009) respectively, *F* (1, 62) < 1.

Distribution of phonemes.

The distribution of phonemes in each of the phoneme positions in the words was balanced as much as possible across the dense and sparse neighborhood density conditions because certain English sounds or sequences of sounds are characteristically difficult for Spanish-accented speakers to produce. For example, for consonants, Spanish-accented speakers tend to produce */z/* as */s/* in the final position, /v/ as /b/ in the initial position, and /p, t, k/ in initial position with less aspiration (Magen, 1998; You, Alwan, Kazemzadeh, & Narayanan, 2005); for vowels, Spanish-accented speakers tend to have more difficulty producing vowels that exist in English but not in Spanish, including /ɪ, æ, ʌ/ (Sidaras et al., 2009). Therefore, these English sounds or sequences of sounds whose production are characteristically difficult for Spanish-accented speakers were all matched among the stimuli in the two neighborhood density conditions.

The distribution of various phonemes among the stimuli in the two neighborhood density conditions is described in the present section. The onset consonants, including /p, t, b, d, f, s, ʃ, n, ɹ / were matched between the dense and sparse neighborhood density conditions. The only unmatched onset consonants were an extra /g/ and /k/ in the sparse neighborhood density condition, and an extra /l/ and /w/ in the dense neighborhood density condition. The vowels, including / i , ɪ, ɝ, e, æ, ɑ, ʌ, ɔ, o, u/ were matched between the dense and sparse neighborhood density conditions. The only unmatched vowels were two extra /ɑu/ in the sparse neighborhood density condition, and two extra /ɑɪ/ in the dense neighborhood density condition.

For the final consonants, / t, d, f, s, ʃ, z, v, ɹ / were matched between the dense and sparse neighborhood density conditions. The unmatched final consonants were those that are not characteristically difficulty for Spanish-accented speakers to produce, including p(4/3), k(3/6), b(2/0), g(2/1), n(5/7), m(3/1), and l(3/4) with their number of occurrence in the sparse and dense conditions in parentheses. Given that there were a certain number of unmatched consonants in the final position, the final consonants were categorized into different manners of articulations (stops, sibilant fricatives, non-sibilant fricatives, nasals, and glides), and its distribution across the dense and sparse conditions was tested. A chi-square test for goodness-of-fit was not significant, $\chi^2 = .128$, $p = .998$, suggesting no statistically significant difference in the distribution of different types of consonants in the final position across conditions.

Overall, the distribution of constituent phonemes in the two conditions was similar; it is more likely that any difference observed in the lexical decision task is due to the difference in the independent variables (*i.e.*, neighborhood density and accent type), rather than difference in the phoneme distribution in the two conditions.

In order to assure the participants were really making lexical decision, a list of 64 phonotactically legal nonwords with the same initial consonant, middle vowel, and phoneme length as the word stimuli were selected from the ARC nonword database as foils to create an equal number of nonword trials (Rastle, Harrington, & Coltheart, 2002). The phonological transcriptions of the nonwords are listed in Appendix B.

Speakers

Two non-native speakers (NNSs) of English (one male and one female) with Spanish as their native language were recruited through flyers sent through the university international student association for recording the spoken word stimuli for the foreign-accented condition. Both speakers were from Lima, Peru and had resided in the U.S. for a minimum of one year but less than two years. The male speaker was 35 years old and learned English when he was 23 years old; the female speaker was 29 years old and learned English when she was 16. Both speakers had learned English after puberty and were judged by 12 native listeners to have a heavy foreign accent in a pilot screening (details are further discussed in the *stimulus preparation* section). Neither speaker reported having hearing or speech disorder. All speakers were paid $10 /hour for their participation.

Two native-speakers (NSs) of American English (one male and one female) from the Midwest were recruited from the University of Kansas to record the native version of the word stimuli under the same conditions as the NNSs.

Recordings

All four speakers (two NNSs and two NSs of English) recorded all the 64 word stimuli and 64 nonword foils. Before the recordings, the NNSs were given the list of stimulus words for practice and then invited to ask for the meaning and pronunciation of any unfamiliar words. To facilitate the recording of the nonwords, the orthographic strings representing the nonwords (e.g., “baith”) were given to the speakers along with the phonologies (e.g., /beɵ/) before and during the recording. A similar sounding real word that was one phoneme different from the nonword to be recorded (e.g. “faith”), as well as its phonology (e.g., /feɵ/), were also provided to the speakers to facilitate recording.

The speakers then practiced reading the whole list aloud for a native English speaker with extensive training in phonetics, who provided assistance with pronunciation for any incorrectly pronounced words. Each of the speakers read each word/ nonword in a random order as presented in the recording list in an IAC sound attenuated booth. In order to generate a few tokens to select from, the speakers were instructed to read three repetitions of each stimulus item. During the recording, the same native English speaker with extensive training in phonetics monitored the whole process and evaluated each production for correctness. Words that were produced incorrectly or too loudly were re-recorded in the same manner. The speech was recorded digitally at a 44.1 kHz sampling rate using a high-quality microphone and a solid-state recorder (Marantz PMD671).

Stimulus Preparation

Each stimulus was edited using Praat (Boersma & Weenink, 2009) into an individual sound file. The amplitude of the individual sound files was increased to their maximum without distorting the sound or changing the pitch of the words by Praat.

Degree of Foreign-accentedness

The degree of foreign-accentedness of each speaker was determined by a foreign-accentedness rating task with 12 native English-speaking pilot listeners from the pool of Introductory Psychology students enrolled at the University of Kansas. The pilot listeners were visually and auditorily presented a random sample of 16 stimulus words (8 dense and 8 sparse words) produced by each of the four speakers in a random order in a noise-free listening condition, and asked to rate each of the 64 items for degree of accentedness using a seven-point scale ranging from 1 (=”native-like”) to 7 (=”strong foreign accent”). Four different versions of presentation were used for counterbalancing purpose such that the 64 stimulus words from each speaker received ratings from three different listeners.

The listeners’ ratings ranged from 1 to 7, suggesting the use of the whole scale. An average rating for each stimulus word produced by each speaker was computed based on the accentedness ratings received. The mean accentedness ratings (standard deviations are in parentheses) for the sparse items, dense items and all items were calculated for each speaker and are listed in Table 1. A 2 (speaker: male vs. female) x 2 (neighborhood density: dense vs. spare) mixed-design ANOVA, with speaker as a within-words factor and neighborhood density as a between-words factor, was conducted on the mean accentedness ratings for the native and accented items respectively. The female native speakers ($M = 1.7$, $SD = .9$) received higher accentedness ratings than the male native speaker ($M = 1.3$, $SD =.6$), $F = 13.39$, $p < .001$. The male ($M = 5.5$, $SD = 1.1$) and female foreign-accented speakers ($M = 5.7$, $SD = 1.0$) both received similar foreign-accentedness ratings. There was no significant difference in accentedness ratings between the sparse and dense items regardless of native-produced or foreign-accented.

An ANOVA with accent type (native vs. foreign-accented) as a within-words factor was conducted on the mean accentedness ratings to check the effectiveness of the accent manipulation. As intended, foreign-accented English speakers ($M = 5.6$, $SD =.9$) were rated as having a stronger accent than were native English speakers ($M = 1.5$, $SD =.6$), $F = 871.4$, $p < .0001$.

Stimulus Duration

The duration of all the stimuli, including word and nonword, were submitted to a 4 (speaker) x 2 (neighborhood density) x 2 (lexicality) mixed-design ANOVAs to check for any speaker effect, neighborhood density effect, lexicality effect, or interaction. The ANOVAs revealed no significant neighborhood density effect (*F* (1, 124) < 1, *p* > .05), nor lexicality effect (*F* (1, 124) < 1, *p* > .05). However, there were a significant speaker effect (*F* (3, 372) = 99.11, *p* < .0001) and a significant interaction between speaker and lexicality (*F* (3, 372) = 6.19, *p* < .0001. The mean word durations (standard deviations are in parentheses) for the word and nonword stimuli were calculated for each speaker and neighborhood density condition, and are listed in Table 2. The mean durations of words and nonwords were significantly different across the speakers. As seen in Table 2, the mean durations of dense and sparse words showed a noticeable difference, ranging from 17 – 36 ms, for the native male and the accented female speakers although the neighborhood density effect was not significant. Given that the reaction time was measured from the onset of the stimuli to the point when the participants responded, thereby it included the duration of the stimuli. These significant differences in stimulus duration across speakers might pose a confounding effect. Therefore, instead of reaction times, corrected reaction times were used during data analysis and are further described in the *result* section.

Counterbalancing Procedure

A 2x2 mixed factorial design that includes *accent type* (and *speaker*) as a between-subjects factor and *neighborhood density* as a within-subjects factor was adopted. To test the listeners’ perceptual adaption to the foreign-accented stimuli, *block* was also included as a within-subjects factor in the experiment. Half (16 items) of the words were randomly selected from each of the two neighborhood density conditions to form list A, and the remaining half formed list B such that each list contained 16 dense and 16 sparse words. For counterbalancing purpose, half of the participants received list A in the first block and list B in the second block (designated by AB in the following paragraph), whereas the other half received list B in the first block and then list A in the second block (designated by BA).

The male and female speakers representing the native-produced and foreign-accented conditions were designated by NMS (native male speaker), NFS (native female speaker), AMS (accented male speaker) and AFS (accented female speaker). With *accent type* and *speaker* as between-subjects factors, participants were randomly assigned to one of the eight counterbalanced conditions, in which all the stimuli were produced only by one of the four speakers (with the order of list presentation in parentheses): NMS-AB, NMS-BA, NFS-AB, NFS-BA, AMS-AB, AMS-BA, AFS-AB, or AFS-BA. Thus, a given listener only heard stimuli spoken by one of the four speakers and each of the 64 stimulus words once – 16 dense and 16 sparse words in the first block and another 16 dense and 16 sparse words in the second block. Items within block were presented in a different randomized order for each participant. Across participants, each stimulus word was presented in both native and foreign-accented form, evenly presented in the two blocks, and evenly represented by each speaker.

Listeners

Forty native speakers (NSs) of American English were recruited from the pool of Introductory Psychology students enrolled at the University of Kansas. The participants received partial credit towards the completion of the course for their participation. All participants were right-handed and reported no history of speech or hearing disorders.

The language experience of all the participants in the current study was surveyed. Since the individual questionnaire was not linked to each participant during data collection, the language experience profile of the participants in each experiment was not available. Instead, the following profile was based on all 160 participants from all four experiments in this study. Around eighty percent of the participants reported having studied Spanish as a second language, but only three of them reported to be fluent in Spanish. Twenty percent of the participants reported to have family members or close friends with a Spanish accent. Three percent of the participants reported having visited or lived in Spanish-speaking countries for more than three months. Around thirty-four percent of the participants have regular contact with non-native speakers of English.

Procedure

Listeners were tested in a group up to three persons each time. Each participant was seated in front of an iMac computer in an individual listening station separated by partitions. The presentation of stimuli and the collection of responses were controlled by PsyScope 1.2.2.

Each trial started with the word “READY” appearing on the computer screen for 500 ms. Then the participants heard one of the randomly selected words or nonwords over a set of Beyerdynamic DT 100 headphones at a comfortable listening level. Each stimulus was presented only once. The participants were instructed to respond as quickly and as accurately as possible whether the item they heard is a real English word or a nonword. If the item is a word, they were to press the button labeled 'WORD' with their right (dominant) hand. If the item is not a word, they were to press the button labeled 'NONWORD' with their left hand. Reaction times were measured from the onset of the stimulus to the onset of the button press response. After the participant pressed a response button, the next trial began. The experiment lasted about 15 minutes. Prior to the experimental trials, each participant received ten practice trials to become familiar with the task. These practice trials were not included in the data analyses.

Results

The current convention in psycholinguistic research is to perform analyses with participants as a random factor (subject analysis, F_1) and with items as a random factor (item analysis, F_2; however see Clark, 1973 for an alternative analysis). However, there is some debate about the proper use and interpretation of additional item analysis over subject analysis, especially when items are carefully matched or balanced across conditions on important variables correlated with the response measures (Raaijmakers, 2003; Raaijmakers, Schrijnemakers, & Gremmen, 1999). Since the stimulus items were well-controlled in the present study, additional item analysis did not seem appropriate or necessary (Raaijmakers et al., 1999). Just to be consistent with the conventions of the field, additional item analyses were reported in all of the experiments in the current study

Reaction times and accuracy rates were the dependent variables of interest. Only accurate responses for the word stimuli were included in the data analysis. Reaction times that were too rapid and too slow (i.e. below 500 ms or above 2000 ms) were considered outliers. Given that the mean duration of word stimuli produced by the fastest speaker (i.e., the native male speaker) in this study was 538 ms, a lower cutoff point of 500 ms was chosen to exclude any responses that were given before the entire stimulus was heard. A more conservative upper cutoff point of 2000 ms was chosen to exclude responses that were obviously out of the boundary of a lexical decision. Using these cutoffs, a total of 1.6% of data, including .78% from the sparse condition and .82% from the dense condition, was excluded from the analysis.

Given that the stimuli spoken by different speakers in the present study had significantly different durations, reaction times were corrected by the stimulus durations in the following analysis as in previous studies (Luce & Pisoni, 1998; Munro & Derwing, 1995b). Raw reaction times were also analyzed with participants as the random variable for reference in Appendix F.1. Since accuracy rates were not influenced by stimulus duration, no adjustment was necessary for accuracy rates. Unless otherwise noted, a significance level of .05 was adopted in all the experiments followed.

With participants as the random variable, responses were pooled across stimulus items, yielding mean reaction times and accuracy rates in the dense and sparse conditions for each participant. To factor out the effect of stimulus duration on the reaction times, the corresponding speaker's mean stimulus duration for the corresponding neighborhood density condition was subtracted from the mean reaction times for each participant, resulting in corrected reaction times. The corrected reaction time is a measurement of the amount of time it takes the participants to press the response button after the end of the utterance. These mean corrected reaction times and accuracy rates for dense and sparse words were then pooled across speakers within in the same accent type condition and subjected to a 2 x 2 mixed-design analysis of variance (ANOVA) with neighborhood density as a within-subjects factor and accent type as a between-subjects factor.

The ANOVA yielded significant main effects of accent type (F_1 (1, 38) = 9.46, p = .004) and neighborhood density (F_1 (1, 38) = 22.95, p < .0001). The mean corrected reaction time for the foreign-accented condition (*M* = 464.46 ms; *SD* = 26.57) is longer than the native-produced condition (*M* = 348.94 ms; *SD* = 26.57). The mean corrected reaction time for the dense words (*M* = 424.68 ms; *SD* = 19.92) is longer than the sparse words (*M* = 388.72 ms; *SD* = 18.37). The interaction between accent type and neighborhood density was also significant, F_1 (1, 38) = 8.56, p =.006. Figure 1 shows the mean corrected reaction times (*S.E.* in parentheses) as a function of accent type (native, accented) and neighborhood density (dense, sparse). The significant interaction was be followed up by simple effect tests with Bonferroni's correction (*p*

< .05). Post hoc tests revealed that dense words were responded to more slowly than sparse words in the accented condition, F_1 (1, 38) = 29.77, $p < .0001$. The same trend was observed in the native condition, but it did not reach statistical significance, F_1 (1, 38) = 1.74, $p = .195$.

It is surprising that the neighborhood density effect was not significant in the native-produced condition given that neighborhood density effect was a robust effect observed in many previous studies using native stimuli (Luce & Pisoni, 1998; Vitevitch, 2003; Vitevitch & Luce, 1998, 1999). It is important to note that most of these previous studies did not correct for stimulus duration given that no significant difference was generally observed in stimulus duration for native-produced dense and sparse words. For the native-produced words in the current study, there was a significant speaker effect observed in the stimulus durations (F (1, 126) = 183.80, $p < .0001$; the native male speaker (M = 535.73 ms, SD = 81.83) produced stimuli with shorter durations than the native female speaker (M = 662.48 ms, SD = 101.08). However, neighborhood density effect (F (1, 126) < 1, $p = .553$) and the interaction between neighborhood density and speaker (F (1, 126) = 1.44, $p = .23$) were not significant on stimulus durations.

To further analyze the subset of data from the native-produced condition in a way more consistent with the previous studies, a 2 (speaker: native male vs. native female) x 2 (neighborhood density: dense vs. spare) mixed-design ANOVA was also conducted on the mean reaction times without correcting for stimulus durations.

Unlike the ANOVA based on corrected reaction times, the present ANOVA revealed a significant neighborhood density effect, F_1 (1, 18) = 5.83, *p* = .027. Native-produced dense words (*M* = 974.44 ms, *SD* = 119.18) were responded to more slowly than sparse words (*M* = 952.43 ms, *SD* = 104.25). The main effect of speaker (*F* (1, 18) = 4.28, *p* = .053) and interaction between speaker and neighborhood density (*F* (1, 18) < 1, *p* = .624) were not significant. This suggests that the set of dense and sparse words used in the present experiment were well-manipulated and replicated the neighborhood density effect consistently found in previous studies provided that the same data analysis procedures were followed.

Accuracy rates were subjected to a 2 x 2 mixed-design analysis of variance (ANOVA) with neighborhood density as a within-subjects factor and accent type as a between-subjects factor. Only the main effects of accent type (F_1 (1, 38) = 36.66, *p* < .0001) and neighborhood density (F_1 (1, 38) = 5.34, *p* = .026) were significant. Participants responded to native-produced words (*M* = 89.00%, *SE*= 1.8%) more accurately than foreign-accented words (*M* = 73.5%, *SE* = 1.8%). In contrast to initial predictions, participants responded to dense words (*M* = 82.34%, *SD* = 11.13%) more accurately than sparse words (*M* = 80.16%, *SD* = 12.00%). Taking together the results from both accuracy rates and corrected reaction times, dense words were responded to more slowly and accurately than sparse words. This might suggest that participants were sacrificing speed for accuracy in making their responses to dense words. However, a significant accent type x neighborhood density interaction that was found in corrected reaction times was not significant in the accuracy rates. This suggests that speed-accuracy trade-off might not be the simple explanation for this result. Another more plausible explanation is explored in the *discussion* section.

To maintain the conventions of the field, item analyses are also reported. With items as the random variable, responses were pooled across subjects within the same speaker group, yielding four sets of mean reaction times and accuracy rates for each item for each speaker group. To factor out the effect of stimulus duration on the reaction times, the stimulus duration (of the corresponding speaker) was subtracted from the mean reaction time for each item, resulting in corrected reaction times (Luce & Pisoni, 1998; Munro & Derwing, 1995b). These four sets of mean corrected reaction times and accuracy rates for each item were then pooled across speakers within in the same accent type condition and subjected to a 2 x 2 mixed-design analysis of variance (ANOVA) with neighborhood density as a between-items factor and accent type as a within-items factor.

For corrected reaction times, the ANOVAs yielded no significant main effects or interaction ($F_2\,(1, 62) = 2.70$, $p = .10$ for the main effect of neighborhood density; all other $F_2 < 1$). For accuracy rates, only the main effect of accent type was significant, $F_2\,(1, 62) = 36.58$, $p < .0001$ (all other $F_2 < 1$). Native-produced words ($M = 88.98\%$, $SD = 11.06\%$) had a higher accuracy rates than the foreign-accented words ($M = 73.52\%$, $SD = 23.87\%$).

To check for any perceptual learning to the foreign-accented stimuli, responses were pooled across stimulus items within the same block, yielding mean corrected reaction times and accuracy rates in each of the two neighborhood density conditions across the two blocks for each participant. A 2 (neighborhood density: dense vs. spare) x 2 (block: first vs. second) repeated-measures analysis of variance (ANOVA), with neighborhood density and block as within-subjects factors, was conducted on the mean corrected reaction times and accuracy rates. Table 3 provides a descriptive summary of the means corrected reaction times (in *ms*) and means accuracy rates (in *percentage*; standard deviations are in parentheses) as a function of neighborhood density (dense vs. sparse) and block of presentation (first vs. second).

For corrected reaction times, participants responded to the second block (M = 429.49, SE= 32.50) more quickly than the first block (M = 500.41, SE= 31.95, F_1 (1, 19) = 14.89, p =.001). The interaction between neighborhood density and block was not significant, F_1 (1, 19) < .01). For accuracy rates, no significant block effect was found. Regardless of neighborhood density, participants showed sign of improvement in their speed in recognizing the foreign-accented words in the second block compared to the first block. This suggests perceptual adaption to the foreign-accented stimuli in the lexical decision task, as reflected only in the response times, and the speed of perceptual learning is not different for dense and sparse words. Since perceptual adaptation to foreign accents is not of central interest in this study and no difference was found in the perceptual adaptation of foreign-accented dense and sparse words, it was not tested in the experiments that followed.

Discussion

Results from Experiment 1 showed that listeners took a longer time to respond to foreign-accented words than native words. This result is consistent with previous studies, which demonstrated that foreign-accented speech takes a longer time to process than native speech (Clarke & Garrett, 2004; Munro & Derwing, 1995b). Results also showed that listeners took a longer time to respond to words from dense than from sparse neighborhoods. More importantly, the significant interaction of accent types and neighborhood density indicated differential effects of neighborhood density as a function of accent type. That is, the native listeners took less time to respond to words from sparse neighborhoods than from dense neighborhoods only in the foreign-accented condition, whereas the same trend did not reach statistical significance in the native-produced condition.

The significant interaction between accent type and neighborhood density in this experiment showed a markedly larger neighborhood density effect for the accented stimuli, relative to the native stimuli. This suggests that lexical discrimination difficulty in dense word recognition is further enhanced in the presence of foreign accent. This result is consistent with the previous results from Imai, et al (2005), which showed an increased processing cost, in term of lower transcription accuracy, for dense words than sparse words in foreign-accented condition relative to native condition. In contrast to Imai, et al (2005), the accent-induced processing cost was not reflected in accuracy rates, but in processing times in the current study. Imai, et al’s study (2005) also differed from the current study in that noise-degraded stimuli were used. Using a new set of well-balanced stimuli without noise-degradation, the current experiment showed that the increased neighborhood density effect on foreign-accented word recognition extends to ideal listening conditions.

Due to substantial pronunciation deviations, accented words are phonologically ambiguous such that it is more difficult to limit the set of possible competing candidates for recognition. When the foreign accents are strong, the listeners tend to rely more heavily on higher-level lexical information to compensate for the inadequate pre-lexical inputs. With so many similar sounding words, there is a higher chance for one or more of these similar sounding words to sound just like the accented pronunciation of the dense target word. Therefore, for foreign-accented dense word, one or more of its similar sounding words might become highly activated and pose a markedly strong competitive effect to the target word. To resolve this increased lexical competition, extra processing time might be needed for the recognition of accented words from a dense neighborhood.

The current accuracy rate result showed that participants responded to native-produced words more accurately than foreign-accented words. This result is consistent with previous studies showing that native listeners transcribed foreign-accented speech with more errors than native speech (Munro & Derwing, 1995a, 1995b). Contrary to prediction, results from accuracy rates also showed that participants responded to dense words more accurately than sparse words, regardless of neighborhood density. Taken together the results from both accuracy rates and reaction times, foreign-accented dense words were responded to more slowly and accurately than sparse words. One of the possibilities is that participants were sacrificing speed for accuracy in making their responses to foreign-accented dense words. However, a significant accent type x neighborhood density interaction that was found in corrected reaction times was not significant in the accuracy rates. This suggests that speed-accuracy trade-off might not be the simple explanation for this result. One possible way to test the hypothesis of speed accuracy trade-off is to run a delayed lexical decision task for the foreign-accented condition to see whether dense words are still responded to more slowly and accurately when participants are given enough time to respond.

Alternatively, the higher accuracy rates for dense words can also be accounted for by the special requirement of the lexical decision task. The lexical decision only requires the participants to decide whether the stimulus item they heard is a real word or not. Therefore, the accuracy rate from the task cannot allow us to assess whether the participants have correctly identify the target word or not. Consider a hypothetic scenario that the participant is given a target word “cat”, and he/she misidentifies the word as “rat” and decides that “rat” is a word. In this case, even though the participant misidentifies the target word, his/ her response is still counted as correct. That means, the accuracy rate in the lexical decision does not really reflect the correct identification of the target words. Especially in the presence of foreign accents, misidentification of the target words as one of its similar sounding words might happen more easily. Substitution of a phoneme might result in another real word more often for dense words than for sparse words. Therefore, the accuracy rates for dense words might be inflated, occurring as an artifact due to the nature of the lexical decision task.

One possible way to test this hypothesis is to ask participants to identify the word that they hear (by typing in or saying out loud the word) after their lexical decision response. This allows us to check whether their accuracy rates for dense words were inflated or not. An alternative way is to run another time-sensitive spoken word recognition task that reflects the correct identification of the word, such as a word naming task.

In sum, using a new set of word stimuli with a new paradigm, Experiment 1 replicated and extended Imai et al's (2005) findings—that the neighborhood density effect was increased on foreign-accented word recognition—to ideal listening conditions. Most importantly, the extra processing cost for foreign-accented dense words was reflected and quantified as longer reaction times in this time-sensitive task that emphasized lexical level of processing. Using reaction time as a measure of processing costs along with tasks that emphasize different levels of processing, Experiments 1 and 2 allow us to induce from which level the accent-induced processing costs stem.

Experiment 2

The primary purpose of Experiments 1 and 2 was to determine the nature of the processing difficulty experienced by native listeners during foreign-accented word recognition. By using tasks that require lexical access for optimal performance, longer reaction times for accented stimuli in Experiment 1 suggests an increased recognition difficulty for accented words rather than native words at the lexical level. This implies that the presence of acoustic-phonetic alternations in foreign-accented speech drive native listeners to increase their use of lexical information during accented word recognition. Given that the stimuli in the lexical decision task underwent both sub-lexical and lexical processing, the reaction times reflected the cumulated processing time from both the sub-lexical and lexica levels. Thus, findings from Experiment 1 did not allow us to imply whether the increased processing difficulty originates from the sub-lexical or lexical level, or both.

There is evidence of foreign accents disrupting sub-lexical processing, such as a higher confusability on vowel recognition in foreign-accented speech than native speech (Sidaras et al., 2009; Wade et al., 2007). It is possible that disrupted sub-lexical processing is the major factor contributing to the increased recognition difficulty for accented words, which is carried over to the lexical level and manifests as longer processing times in the lexical-level-emphasizing task. The present experiment sought to further examine the level of processing through which foreign accents induce recognition difficulties. Specifically, we attempted to investigate whether the processing difficulty of foreign-accented speech, as demonstrated by increased reaction times in Experiment 1, originates from the sub-lexical or lexical level, or both.

To examine whether the processing difficulty of foreign-accented stimuli found in Experiment 1 actually stems from the sub-lexical level or not, the current experiment used the same set of stimuli from Experiment 1, but with a task that emphasizes the sub-lexical level of processing— a speeded same-or-different-matching task (a.k.a., AX task). Through biasing the listeners to process the same subset of real words using sub-lexical representations in one task (AX task in the current experiment), and using lexical representations in another task (lexical decision task in Experiment 1), we could compare the processing difficulty across tasks, and distinguish the processing costs originating from the sub-lexical and lexical levels respectively. Therefore, we could determine whether the sub-lexical or lexical level of processing poses greater recognition difficulties for foreign-accented words.

In the AX task, the participants were presented with two spoken stimuli in a row and required to respond as quickly and as accurately as possible if the two items are the same or different. As this experimental task only requires low-level matching of two acoustic patterns, lexical activation and the lexical level of processing involved is assumed to be minimal. In an attempt to further bias listeners to use sub-lexical representations to process the spoken stimuli in this task, a significantly greater proportion of nonword pairs compared to word pairs (a ratio of 3:1) were used as stimuli. Moreover, instead of presenting the words and nonwords in separate blocks, the two sets of stimuli were intermixed during the presentation. These strategies, adopted from Vitevitch and Luce (1999) and Vitevitch (2003), have been shown to be effective to promote sub-lexical processing of both the words and nonwords in this task.

To allow a direct comparison of the reaction times across experiments, the word stimuli varying in neighborhood density from Experiment 1 were used as stimuli in the AX task in the current experiment. As in Experiment 1, the stimuli were spoken by native and foreign-accented speakers to form the native-produced and foreign-accented conditions. Listeners' response times and accuracy rates as a function of neighborhood density, accent type and task (the current AX task vs. lexical decision task in Experiment 2) were of interest.

Of particular interest to the current experiment was the locus of difficulties in recognizing foreign-accented words varying in neighborhood density. The acoustic-phonetic deviations in accented speech signals might drive listeners to rely more heavily on lexical information for word recognition. This might especially enhance the competitive effects posed by the dense neighborhoods, leading to an increased difficulty for dense words. Thus, dense words yielded longer reaction times than sparse words, especially in the foreign-accented condition in the lexical decision task in Experiment 1. This extra processing cost for foreign-accented words may arise from the lexical or sub-lexical level, or both.

If the increased processing difficulty for dense words in the accented condition mainly arises from the sub-lexical level, we would expect the results from the present AX task, which emphasizes sub-lexical level of processing, to resemble those from the lexical decision task in Experiment 1, which emphasizes the lexical level of processing. Otherwise, if the increased processing difficulty for dense words mainly arises from the lexical level rather than the sub-lexical level, we would see a markedly different pattern of results from the two tasks. There could be many different possible outcomes. For example, similar reaction times may be observed for dense and sparse words regardless of accent type, suggesting that foreign accents do not impact sub-lexical processing at all. It is also possible that increased reaction times are observed for accented words relative to native words, but the neighborhood density effect is not significant. This scenario implies that foreign accents impair sub-lexical processing in general, but the locus of accent-induced processing difficulty for dense words is at the lexical level.

Method

Stimuli and design

The same 32 dense and 32 sparse words from Experiment 1 were used as stimuli to serve as SAME pairs in the present experiment. Noise-free sound recordings of these 64 word stimuli from the two native and two foreign-accented speakers used in Experiment 1 were used in the present experiment. In order to promote sub-lexical processing of words in this task, 192 English nonwords were used as foils to create a significantly greater nonword to word proportion (3: 1). These nonwords were monosyllabic and consisting of three phonemes in a consonant-vowel-consonant structure. They were all phonologically legal syllables of English selected from the ARC nonword database (Rastle et al., 2002) and are listed in Appendix C.

To assure that the participants were really discriminating the stimulus pairs rather than responding ‘SAME’ all the time, an equal number of filler items served as DIFFERENT pairs. Two hundred and fifty-six nonwords with the same onset consonant, middle vowel, and phoneme length as the (SAME pair) stimuli were used as filler items to form 256 DIFFERENT pairs. These nonword foils for forming the DIFFERENT pairs are listed in Appendix D. Responses to the nonword SAME foil pairs, as well as the DIFFERENT filler pairs, were not be included in the data analysis.

Speakers and Recordings

All the stimuli used in this experiment were recorded by the same native and foreign-accented speakers in the same manner and at the same time as the other stimuli that were used in Experiment 1.

Stimulus Preparation

The recording of each stimulus was prepared in the same way as in Experiment 1 using Praat (Boersma & Weenink, 2009), such as normalization and intelligibility piloting.

Counterbalancing Procedure

The male and female speakers representing the native-produced and foreign-accented conditions were designated by NMS (native male speaker), NFS (native female speaker), AMS (accented male speaker) and AFS (accented female speaker). With accent type and speaker as between-subjects factors and neighborhood density as a within-subjects factor, participants were randomly assigned to one of the four counterbalanced conditions, in which all the stimuli were produced only by one of the four speakers: NMS, NFS, AMS, or AFS. The order of stimulus presentation within each listener was randomized regardless of neighborhood density and stimulus type (stimulus or filler). Each listener received 512 trials and heard each of the 64 word stimulus SAME pairs, 192 nonword-foil SAME pairs, and the corresponding 256 DIFFERENT filler pairs only once, in a different randomized presentation order.

Listeners

Forty native speakers (NS) of American English were recruited from the pool of Introductory Psychology students enrolled at the University of Kansas. The participants received partial credit towards the completion of the course for their participation. All participants were right-handed and reported no history of speech or hearing disorders. None of the participants in the present experiment took part in any of the other experiments in this study.

Procedure

Listeners were tested in a group of up to three persons each time. Each participant seated in front of an iMac computer in an individual listening station separated by partitions. PsyScope 1.2.2 were used to control the randomization and presentation of stimuli. A New Micros response box that contains a dedicated timing board was connected to the iMac computer to provide millisecond accuracy for response collection. In each trial, the word "READY" appeared on the computer screen for 500 ms. Participants then heard one pair of the randomly selected stimuli or fillers through a set of Beyerdynamic DT 100 headphones at a comfortable listening level. A 50 ms interstimulus interval was used. The participants were instructed to respond as quickly and as accurately as possible whether the two items they hear are the SAME or DIFFERENT. If the items are the SAME, they are to press the button labeled ‘SAME’ with the right (dominant) hand. If the items are DIFFERENT, they are to press the button labeled ‘DIFFERENT’ with their left hand. Reaction times were measured from the onset of the second stimulus in the pair to the button press response. After the participant pressed the response button, the next trial began. Every participant received a total of 512 trials. Half of the stimulus pairs were the SAME pairs, and half of the stimulus pairs were the DIFFERENT filler items. The experiment lasted about 30 minutes. Prior to the experimental trials, each participant received ten practice trials to become familiar with the task. These practice trials were not included in the data analyses.

Results

Reaction times and accuracy rates were the dependent variables of interest. Responses to the foils and fillers were not included in the data analysis. Only accurate responses to the SAME word pairs were included. To allow a direct comparison of results from the current experiment and Experiment 1, the same cutoff criteria were used to exclude outliners. That is, reaction times that are too rapid or too slow (i.e. below 500 ms and above 2000 ms) were considered outliners and excluded from the analysis. A total of 4.4% of data, including 2.1% from the sparse condition and 2.3% from the dense condition, was excluded from the analysis. Reaction times in the current experiment were adjusted by stimulus durations as in Experiment 1 to result in corrected reaction times. Raw reaction times were also analyzed with participants as the random variable for reference in Appendix F.2.

With participants as the random variable, corrected reaction times and accuracy rates from the LD (lexical decision task; Experiment 2) and the current AX task were subjected to 2 x 2 x2 mixed-design ANOVAs with neighborhood density as a within-subjects factor, and accent type and task as between-subjects factors. Only significant results that are relevant and of interest are reported here. Significant interactions were followed up by simple effect tests with Bonferroni's correction ($p < .05$).

For corrected reaction times, the main effects of neighborhood density effect (F_1 (1, 76) = 11.13, p = .001) and task (F_1 (1, 76) = 96.60, $p < .0001$) were significant. Dense words (M = 297.17, SE = 12.27) were responded to more slowly than sparse words (M = 279.04, SE = 12.46). Participants took a longer time to respond to the LD task (M = 406.70, SE = 17.06) than the AX task (M = 169.51, SE = 17.06). This result that responses in the AX task were shorter than those in the LD task is consistent with our assumption that the stimuli in the AX task only underwent sub-lexical processing, whereas the stimuli in the LD task underwent both sub-lexical and lexical processing.

There was a significant two-way interaction between accent type and task, F_1 (1, 76) = 11.31, p = .001. Post hoc tests indicated that foreign-accented words were responded to more slowly than native-produced words in the LD task (F_1 (1, 76) = 11.46, p = .0001), whereas foreign accents did not induce any processing cost on reaction times in the AX task (F_1 (1, 76) = 1.88, p = .17). The means corrected reaction times (*S.E.* in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks are presented in Figure 2a. The two-way interaction between neighborhood density and task was also significant, F_1 (1, 76) = 8.75, p = .004. Post hoc tests indicated that dense words were responded to more slowly than sparse words in the LD task (F_1 (1, 76) = 21.91, $p < .0001$), whereas no significant neighborhood density effect was observed in the AX task ($F < 1$). The mean corrected reaction times (*S.E.* in parentheses) for the dense and sparse words in the AX and LD tasks are presented in Figure 3. The three-way interaction was not significant, F_1 (1, 76) = 1.17, p = .28.

For accuracy rates, the ANOVA revealed significant main effects of accent type (F_1 (1, 76) = 21.71, $p < .0001$) and task (F_1 (1, 76) = 64.20, $p < .0001$). Participants responded to native-produced words (M = 91.1%, SE = .011) more accurately than accented words (M = 83.9%, SE = .011). Participants attained higher accuracy in the AX task (M = 93.8%, SE = .011) than the LD task (M = 81.2%, SE = .011), suggesting that sub-lexical processing of the word stimuli is less challenging than the lexical processing for the native listeners. This is not surprising given that the sub-lexical processing in the AX task only required the listeners to match two low-level acoustic patterns to accurately discriminate the two words.

There was also a significant two-way interaction between accent type and task, F_1 (1, 76) = 27.67, $p < .0001$. Post hoc tests indicated that foreign-accented words were responded less accurately than native-produced words in the LD task (F_1 (1, 76) = 49.20, $p < .0001$), whereas foreign accent did not induce any processing cost on accuracy rates in the AX task ($F_1 < 1$). The means accuracy rates (*S.E.* in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks are presented in Figure 4a. All other main effect and interactions were not significant.

With items as the random variable, corrected reaction times and accuracy rates from the LD (lexical decision task; Experiment 2) and the current AX task were subjected to 2 x 2 x2 mixed-design ANOVAs with neighborhood density as a between-words factor and accent type and task as within-words factors. Significant interactions were be followed up by simple effect tests with Bonferroni's correction ($p < .05$).

For corrected reaction times, the ANOVA yielded significant main effects of accent type (F_2 (1, 62) = 13.68, $p < .0001$) and task (F_2 (1, 62) = 488.41, $p < .0001$). Corrected reaction times were longer for native-produced (*M* = 285.40, *SE* = 9.28) than foreign-accented words (*M* = 263.27, *SE* = 8.55). Corrected reaction times for the AX task (*M* = 168.56, *SE* = 7.95) were shorter than the LD task (*M* = 380.12, *SE* = 11.14). Only the two-way interaction between accent type and task was also significant, F_2 (1, 62) = 10.23, $p = .002$. Post hoc tests indicated that foreign-accented words were responded more quickly than native-produced words in the AX task (*F* (1, 62) = 28.80, $p < .0001$), whereas such processing advantage for foreign-accented words were not found in the LD task ($F < 1$). The means corrected reaction times (*S.E.* in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks are presented in Figure 2b.

For accuracy rates, the ANOVA revealed significant main effects of accent type, F_2 (1, 62) =27.95, $p < .0001$, and task, F_2 (1, 62) = 37.44, $p < .0001$. Participants responded to native-produced words (*M* = 91.1%, *SE* = .008) more accurately than accented words (*M* = 83.9%, *SE* = .016). Participants attained higher accuracy in the AX task (*M* = 93.8%, *SE* = .006) than the LD task (*M* = 81.3%, *SE* = .020). There was also a significant two-way interaction between accent type and task, F_2 (1, 62) = 37.41, $p < .0001$. Post hoc tests indicated that foreign-accented words were responded less accurately than native-produced words in the LD task (F_2 (1, 62) = 36.58, $p < .0001$), whereas foreign accent did not induce any processing cost on accuracy rates in the AX task (F_2 = 1.06, $p = .308$). The means accuracy rates (*S.E.* in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks are presented in Figure 4b. All other main effect and interactions were not significant.

Discussion

The current experiment aimed to determine the locus of foreign-accented word recognition difficulty by comparing the reaction times across the AX and LD tasks. The overall reaction times for the AX task were significantly shorter than those for the LD task. This is consistent with the assumption that the stimuli in the current AX task undergo only sub-lexical processing with no or minimal lexical activation, whereas the stimuli in the LD task undergo both sub-lexical and lexical processing. Without the extra level of lexical processing in the AX task, participants took a shorter time to respond to the AX task relative to the LD task. Overall, participants also attained higher accuracies in the AX task than the LD task. This suggests that sub-lexical processing in the AX task were easier for the participants than the lexical processing in the LD task.

Although foreign accent induced processing costs, including longer reaction times and lower accuracy rates, in the LD task, participants responded to the foreign-accented and native-produced stimuli similarly, in terms of both reaction times and accuracy rates, in the AX task. The absence of accent type effect in the AX task suggests that the sub-lexical processing of the foreign-accented words is as good as that of the native-produced words in the AX task. It seems that foreign accents only impair the lexical level of processing, but have no influence on the sub-lexical level of processing, at least in the current AX task. It is important to acknowledge the inherent difficulty in interpreting null-results in the current AX task, thus these results should be considered with caution. However, considering that participants attained a high average accuracy rate of 93.8% in the AX task, the task should be sensitive enough to detect any differences if they exist.

The finding that dense words were responded to more slowly than sparse words only in the LD task but not in the AX task suggests that the processing difficulties for dense words in both the native and accented conditions in the LD task mainly originate from the lexical level. In the LD task, words were process lexically such that lexical representations of both the target word and its similar sounding words were activated. Hence, there was a strong lexical competition, especially for target words with a dense neighborhood, resulting in longer reaction time for dense words in the LD task. In contrast, similar sounding words from the phonological neighborhoods were not activated to exert any competitive effect on the dense target words in the AX task. This implies that only sub-lexical or acoustic-phonetic representations of the words were activated in the AX task regardless of accent type. Despite the greater acoustic-phonetic deviations and variability in the foreign-accented stimuli, sub-lexical processing was sufficient for efficient performance in the AX task.

The advantage for sub-lexical processing of foreign-accented words over native-produced words observed in the item analyses might seem surprising. However, it might simply suggest that different strategies were adopted for the sub-lexical processing of foreign-accented and native-produced words. It is possible that the native listeners were just simply doing low-level acoustic matching for foreign-accented words. That is, the distortion of a foreign accent resulted in listeners treating the foreign-accented words much like they treat environmental sounds, using basic perceptual processes to respond in the AX task. Instead for the native words, the listeners activate the corresponding sub-lexical (and perhaps to some extent lexical) representations in the matching process, much like readers in the Stroop task cannot avoid reading the word instead of the ink color (Stroop, 1935), so that a longer time was needed for the sub-lexical processing of the native words.

The idea that listeners are compelled to process and activate words when they hear their native speech was evidenced in previous research, which compared the masking effects of native and non-native interfering (background) speech on the recognition of a target speech. Several studies consistently showed that background speech in listeners' native language is more distracting than background speech in a foreign language when the listeners were asked to identify target speech in their native language (Rhebergen, Versfeld, & Dreschler, 2005; Van Engen, 2010; Van Engen & Bradlow, 2007). Transcriptions by participants often contained intrusions from the background native speech, demonstrating listeners' difficulty in tuning out the native background speech (Van Engen & Bradlow, 2007). Thus, considering these previous findings, it is not surprising that the foreign-accented and native-produced words were processed differently in the current AX task. It is important to keep in mind that the results from the item analyses might not be appropriate given the well-controlled stimulus set and have to be considered with caution.

Overall, the cross-task analysis suggests that despite the substantial pronunciation deviations in the foreign-accented words, the listeners managed to make same-difference discrimination efficiently in the AX task without increasing their use of lexical information. Therefore, the competing words from the dense neighborhood are not activated to exert an effect on the sub-lexical processing of accented dense target words in the AX task. These results imply that the extra processing cost for foreign-accented dense words shown in the lexical decision task mainly originated from the lexical level.

Experiment 3

Thus far, Experiments 1 -2 focused on the influence of foreign accents on the recognition of words varying in neighborhood density. As neighborhood density is a lexical factor, Experiments 1-2 allowed us to examine the impact of foreign accents on the lexical level of processing. Results from Experiments 1 & 2 suggest that the processing costs associated with foreign accents mainly originate from the lexical level. However, as mentioned earlier, the gross phonetic and phonological deviations induced by foreign accents were shown to disrupt sub-lexical processing, resulting in a higher confusability in vowel recognition of foreign-accented speech (Sidaras et al., 2009; Wade et al., 2007). Therefore, the impact of foreign accents on the sub-lexical level of processing was further investigated in Experiment 3 using word and nonword stimuli varying in a sub-lexical factor.

The primary purpose of Experiment 3 was to test whether foreign accents impact the processing of word-like segments varying in the sub-lexical characteristic— phonotactic probability. Phonotactic probability is a sub-lexical frequency referring to the relative frequencies of segments and sequences of segments in syllables and words (Vitevitch & Luce, 1999). This sub-lexical factor is of particular interest here, as it has been demonstrated to facilitate native-produced word recognition at the sub-lexical level (Vitevitch, 2003; Vitevitch & Luce, 1998, 1999). Given the distortion of sub-lexical inputs found in foreign-accented speech (Sidaras et al., 2009; Wade et al., 2007), phonotactic probability may play an important role in processing. It is hypothesized that native listeners may rely more heavily on phonotactic probability in the sub-lexical processing of foreign-accented stimuli compared to native-produced stimuli. Thus, accented stimuli that contain common segments and sequences of segments (i.e., high phonotactic probability) are expected to be processed more effectively than those that contain less common segments and sequences of segments (i.e., low phonotactic probability) by native listeners, whereas the native stimuli are expected to show a relatively smaller phonotactic probability effect.

A speeded AX task that emphasizes the sub-lexical level of processing, as in Experiment 2, was used in the present experiment. As in Experiment 2, a significantly greater proportion of nonword pairs compared to word pairs were used as stimuli to further bias listeners to use sub-lexical representations to process the spoken stimuli in this task. Unlike Experiment 2, the nonword pairs did not serve as foils, but some of them served as stimuli along with the word pairs. Both the words and nonwords were manipulated on phonotactic probability to form the high and low phonotactic probability conditions, which were controlled on neighborhood density. All the stimuli were spoken by native and foreign-accented speakers to form the native-produced and foreign-accented conditions. Listeners' response times and accuracy rates to the SAME responses as a function of phonotactic probability and accent types were of interest.

Regardless of lexicality, native listeners were expected to respond SAME to high phonotactic probability items faster and more accurately than to low phonotactic probability items, especially for foreign-accented stimuli. A neighborhood density effect was not expected for the word stimuli as they were expected to undergo sub-lexical level of processing, and also the two phonotactic probability conditions were controlled on neighborhood density.

Method

Stimuli and design

A hundred and ninety-two nonwords and fifty-six words in English that are monosyllabic and consisting of three phonemes in a consonant-vowel-consonant structure were used as stimuli to serve as SAME pairs in the present experiment. The nonwords are all phonologically legal syllables of English selected from the ARC nonword database (Rastle et al., 2002). All the 56 words and 72 of the nonwords were carefully selected as stimuli to form the high and low phonotactic probability sets, resulting in 28 words and 36 nonwords in each set. The phonotactic probability was measured by (1) how often a certain segment occurs in a certain position in a word (positional segment frequency) and (2) the segment-to-segment co-occurrence probability (biphone frequency; Vitevitch & Luce, 1998).

The high and low phonotactic probability word sets (listed in Appendix E.1-E.2) were controlled on subjective familiarity, mean log word frequency, raw word frequency, neighborhood density, and mean log neighborhood frequency. The sub-lexical and lexical characteristics of the high and low phonotactic probability word lists are also presented in Appendix E.1-E.2, whereas the descriptive statistics and ANOVAs of the mean values are presented in Table 4.

The high and low phonotactic probability nonword sets (listed in Appendix E.3-E.4) were controlled on neighborhood density and mean log neighborhood frequency. Descriptive statistics for the sub-lexical and lexical characteristics of the high and low phonotactic probability nonword lists and ANOVAs of the mean values are presented in Table 5.

The rest of the 120 nonwords were used as foils to create a significantly greater nonword to word proportion (3.4 : 1) to promote sub-lexical processing of words in this task. These nonword foils that served as SAME pairs are listed in Appendix E.5. In order to assure that the participants were really discriminating the stimulus pairs rather than responding ‘SAME’ all the time, an equal number of filler pairs served as DIFFERENT pairs. Two hundred and forty-eight nonwords with a different final phoneme as the stimuli and foils were chosen to pair up with the stimuli and foils to form the DIFFERENT pairs. For example, a SAME pair, ‘bag bag’ has one corresponding DIFFERENT pair, ‘bag bab.’ The 248 nonwords for forming the DIFFERENT pairs are listed in Appendix E.6.

Distribution of phonemes.

The distribution of phonemes in each of the phoneme positions in the words was balanced as much as possible across the high and low phonotactic probability conditions, especially for those English sounds or sequences of sounds whose production is characteristically difficulty for Spanish-accented speakers. The onset consonants, including /p, t, d, f, l, ɹ / were matched between the high and low phonotactic probability conditions. The unmatched onset consonants (the number of unmatched occurrence is in parentheses) were extra /g (1), n (2), ɹ (1), w (2)/ in the low phonotactic probability condition, and extra /b (1), k (2), m (1) and s (2)/ in the high phonotactic probability condition.

The vowels, including /ɪ, ɝ, o, u/ were matched between the two phonotactic probability conditions. Spanish and English only share five vowels, including /i, e, ɑ, o, u/. English vowels that do not exist in the Spanish phonemic inventory might be particularly difficult for Spanish-native speakers to produce. Hence the number of unmatched vowels that do not exist in the Spanish phonemic inventory was balanced across the two conditions. The unmatched vowels were /e (1), ɑɪ (1), ʌ (1), ʊ (1), ɝ (1)/ in the low phonotactic probability condition, and /ɪ (2), æ (2), ɑ (1)/ in the high phonotactic probability condition. Each condition contained five unmatched vowels—four that do not exist in the Spanish phonemic inventory and one that exist in both inventories.

For the final consonants, / d, θ, ʃ, dʒ, tʃ, m, k, z, v / were matched between the two phonotactic probability conditions. The unmatched final consonants were those that are not characteristically difficulty for Spanish-accented speakers to produce in the final position, including p(0/1), f(0/1), b(0/1), g(0/1), n(1/0), s(1/0), and l(2/0) with their number of unbalanced occurrence in the high and low phonotactic probability conditions in parentheses.

Likewise, the distribution of phonemes in each of the phoneme positions in the nonwords was balanced as much as possible across the high and low phonotactic probability conditions. The onset consonants, including / t, b, d, tʃ, dʒ, l, n, ʃ, θ, l, ɹ, v / were matched between the high and low phonotactic probability conditions. The unmatched onset consonants (the number of unbalanced occurrence is in parentheses) were /f (1), g (1), n (1), j (1), z (1)/ in the low phonotactic probability condition, and / k (1), m (1) and s (3)/ in the high phonotactic probability condition. The vowels, including /i, ɪ, o, ɑɪ, ʌ / were matched between the two phonotactic probability conditions. The unmatched vowels were /aʊ (2), ɝ (1), u (1)/ in the low phonotactic probability condition, and /ɪ (1), æ (1), ɑ (1), ɛ (1)/ in the high phonotactic probability condition. For the final consonants, / t, θ, ʃ, dʒ, tʃ, m, z, v / were matched between the two phonotactic probability conditions. The unmatched final consonants were /f (2), g (2), p (1), b (1)/ in the low phonotactic probability condition, and /k (1), s (3), l (1), n (1)/ in the high phonotactic probability condition.

Speakers and Recordings

The stimuli used in this experiment were recorded by the same native and foreign-accented speakers in the same manner and at the same time as the other stimuli that were used in Experiments 1 and 2.

Duration

The duration of all the 56 word and 72 nonword stimuli for each speaker were submitted to a 4 (speaker) x 2 (phonotactic probability) x 2 (lexicality) mixed-design

ANOVAs to check for any speaker effect, phonotactic probability effect, lexicality effect, or interaction. The ANOVAs revealed no significant phonotactic probability effect, F (1, 124) = 1.47, p= .227. However, there were a significant speaker effect (F (3, 372) = 118.60, p < .0001), lexicality effect (F (1, 124) = 29.22, p < .0001), and a significant interaction between speaker and lexicality (F (3, 372) = 11.73, p < .0001). The mean word durations (standard deviations are in parentheses) for the word and nonword stimuli for each speaker and phonotactic probability condition are listed in Table 6. The durations of words and nonwords were different across the speakers such that stimulus duration might pose a confounding effect. As in Experiments 1 and 2, corrected reaction times were used to eliminate this possible confounding during data analysis (Luce & Pisoni, 1998; Munro & Derwing, 1995b).

Stimulus Preparation

The recording of each stimulus was prepared in the same way as in Experiments 1 -2 using Praat (Boersma & Weenink, 2009), such as normalization and intelligibility piloting.

Counterbalancing Procedure

Presentation of the nonword and word stimuli and their corresponding filler pairs followed the same counterbalancing procedure used in Experiment 2. With *accent type* and *speaker* as between-subjects factors and phonotactic probability as a within-subjects factor, participants were randomly assigned to one of the four counterbalanced conditions, NMS, NFS, AMS, or AFS, in which all the stimuli were produced only by one of the four speakers. The order of stimulus presentation within each listener was randomized regardless of phonotactic probability and stimulus type (stimulus or filler). Each listener received 496 trials and heard each of the 56 word stimulus SAME pairs, 72 nonword stimulus SAME pairs, 120 nonword-foil SAME pairs, and the corresponding 248 DIFFERENT filler pairs only once, in a different randomized presentation order.

Listeners

Forty native speakers (NS) of American English were recruited from the pool of Introductory Psychology students enrolled at the University of Kansas. The participants received partial credit towards the completion of the course for their participation. All participants were right-handed and reported no history of speech or hearing disorders. None of the participants in the present experiment took part in any of the other experiments in this study.

Procedure

The procedure was the same as those in Experiment 2.

Results

Reaction times and accuracy rates were the dependent variables of interest. Only accurate responses for the 56 word and 72 nonword stimulus SAME pairs were included in the analysis. Like Experiments 1 and 2, reaction times that were too rapid or too slow (i.e. below 500 ms or above 2000 ms) were considered to be outliers and excluded from the analysis. For the word condition, a total of 5.8% of data, including 2.8% from the low phonotactic probability condition and 3% from the high phonotactic probability condition, was excluded from the analysis. For the nonword condition, a total of 3.4% of data, including 1.7% from each of the phonotactic probability conditions, was excluded from the analysis. Reaction times in the current experiment were adjusted by stimulus durations as in Experiment 1 to result in corrected reaction times. Raw reaction times were also analyzed with participants as the random variable for reference in Appendix F.3.

With participants as the random variable, corrected reaction times and accuracy rates were subjected to 2 x 2 x2 mixed-design ANOVAs with phonotactic probability and lexicality as within-subjects factors, and accent type as a between-subjects factor. For corrected reaction times, the ANOVA yielded a significant main effect of lexicality, F_1 (1, 38) = 10.10, $p = .003$, and a significant two-way interaction between lexicality and accent type, F_1 (1, 38) = 14.04, $p = .001$. The significant interaction was followed up by simple effects tests with Bonferroni's correction ($p < .05$). Post hoc analysis showed that participants took longer to respond to words than nonwords in the foreign-accented condition, F (1, 38) = 23.98, $p < .0001$, whereas no such difference was found in the native-produced condition, $F < 1$. The means corrected reaction times (*S.E.* in parentheses) for the words and nonwords in the native-produced and foreign-accented conditions are presented in Figure 5a. All other main effects and interactions were not significant (all $F < 1$).

For accuracy rates, the ANOVA revealed a marginally significant lexicality effect, *F* (1, 38) = 4.09, *p* = .05, and a significant two-way interaction between lexicality and accent type, *F* (1, 38) = 9.76, *p* = .003. Post hoc analysis showed that words had a lower accuracy rates than nonwords in the foreign-accented condition, *F* (1, 38) = 13.24, *p* = .001, whereas no such difference was found in the native-produced condition, *F* < 1. The mean accuracy rates (S.E. in parentheses) for the words and nonwords in the foreign-accented and native-produced conditions are presented in Figure 6a. All other main effects and interactions were not significant.

With items as the random variable, corrected reaction times and accuracy rates were subjected to 2 x 2 x2 mixed-design ANOVAs with phonotactic probability and lexicality as between-words factors and accent type as a within-words factor. Significant interactions were be followed up by simple effect tests with Bonferroni's correction (*p* < .05). For corrected reaction times, the ANOVA yielded a significant main effect of lexicality, F_2 (1, 124) = 4.54.10, *p* = .035, and a significant two-way interaction between lexicality and accent type, F_2 (1, 124) = 13.17, *p* <.0001. The significant interaction was followed up by simple effects tests with Bonferroni's correction (*p* < .05). Post hoc analysis showed that participants took longer to respond to words than nonwords in the foreign-accented condition, *F* (1, 124) = 11.23, *p* = .001, whereas no such difference was found in the native-produced condition, *F* < 1. The means corrected reaction times (*S.E.* in parentheses) for the words and nonwords in the native-produced and foreign-accented conditions are presented in Figure 5b. All other main effects and interactions were not significant.

For accuracy rates, the ANOVA revealed a significant two-way interaction between accent type and lexicality, F_2 (1, 124) = 11.01, $p = .001$. Post hoc analysis showed that words had a lower accuracy rates than nonwords in the foreign-accented condition, F (1, 124) = 11.86, $p = .001$, whereas no such difference was found in the native-produced condition, $F < 1$. The mean accuracy rates (S.E. in parentheses) for the words and nonwords in the foreign-accented and native-produced conditions are presented in Figure 6b. Post hoc analysis also showed that foreign accent reduced accuracy only when the items were words, F (1, 124) = 12.04, $p = .001$, whereas no such difference was found for nonwords, F (1, 124) = 1.17, $p = .28$. The mean accuracy rates (S.E. in parentheses) for the native-produce and foreign-accented items in the two lexicality conditions are presented in Figure 7. All other main effects and interactions were not significant.

The current task was designed to promote the use of sub-lexical representations to process both the real words and nonwords by presenting mostly nonword stimuli to the participants. Thus, both the nonword and word stimuli were expected to be processed sub-lexically, and lead to similar patterns of results in the current task. However, longer corrected reaction times and lower accuracy rates were found for words than nonwords in the foreign-accented condition, but not in the native-produced condition. With the distortions introduced by foreign accents, processing of words in the AX task was subjected to higher processing costs over nonwords. This significant interaction of accent type and lexicality indicates differential effects of lexicality as a function of accent type. This result suggests that the words and nonwords were processed differently only in the presence of foreign accents.

There are two possibilities that can account for the current result. First, it is possible that the acoustic-phonetic alternations induced by foreign accents were severe that listeners experienced increased difficulty in sub-lexical processing of foreign-accented words, which they compensate for by using top-down lexical knowledge. With the availability of top-down lexical knowledge, the words were processed lexically, whereas the nonwords were processed sub-lexically. Therefore, more time is required to activate the additional lexical information to help with the sub-lexical processing words in the AX task. Second, in the presence of foreign accent, it might be more efficient for the listeners to process the nonword stimuli using only low-level acoustic matching strategy to make the same-different discrimination. Instead, the word stimuli were still processed sub-lexically. Therefore, more time and effort was required to sub-lexically process the words relative than the nonwords in the AX task.

To distinguish these two possible explanations, post hoc tests were conducted on the corrected reaction times and accuracy rates for the native-produced and foreign-accented words and nonwords respectively. However, no significant differences were found on either corrected reaction times (F_1 (1, 38) = .50, p = .48) or accuracy rates (F_1 (1, 38) = 1.49, p = .23) for the native-produced and foreign-accented words. Similarly, no significant differences were found on either corrected reaction times (F_1 (1, 38) = 1.24, p = .27) or accuracy rates (F_1 (1, 38) = .21, p = .65) for the native-produced and foreign-accented nonwords. Thus, based on the post hoc tests, it is not clear whether the words were processed lexically or the nonwords were processed using low-lexical matching strategy in the presence of foreign accents.

Remember from Experiments 1 and 2 that both the accent type and the neighborhood density effects, which were significant in the LD task, were not significant in the AX task. These results suggest that the sub-lexical processing of the word stimuli in the AX task were not negatively impacted by foreign accents. Given that the AX task in the current experiment mirrored that in Experiment 2, it is reasonable to consider the current result in the context of those from Experiment 2. Taken together the results from the current experiment and Experiment 2, it seems more plausible that the nonwords were processed using a low-level matching strategy, rather than the words were processed lexically in the AX task.

Discussion

Neither the main effect of phonotactic probability nor its interaction of any kind was significant. This result is surprising given that phonotactic probability has been demonstrated to facilitate word recognition at the sub-lexical level in native speech (Vitevitch, 2003; Vitevitch & Luce, 1998, 1999). The absence of phonotactic probability effect even in the native-produced condition suggests that the phonotactic probability effect was not strong enough for detection in the current experiment. It is important to note the differences in the stimulus set used in the current experiment and the previous studies. To minimize the possible influence from phonemes whose production is characteristically difficult for Spanish-accented speakers, the distribution of phonemes in each of the phoneme positions in the stimuli was balanced as much as possible across the two phonotactic probability conditions in the current experiment.

Due to this extreme balancing measure, selecting stimuli with suitable properties for this experiment was very difficult. The manipulation of phonotactic probability, including both segment and biphone probabilities, was much smaller relative to the previous studies (Vitevitch and Luce, 1999; Vitevitch, 2003). Table 7 presents the average segment and biphone probability for the word and nonword stimuli in the high and low phonotactic probability conditions and the corresponding magnitude differences between the two conditions in Vitevitch and Luce (1999), Vitevitch (2003) and the current experiment. Hence, with a strict balancing measure, the two phonotactic probability conditions in the current experiment might not be different enough to generate a detectable phonotactic probability effect. It is not clear whether the absence of phonotactic probability effect in the foreign-accented condition is because the effect is not strong enough to be detected or the effect manifested the same way in foreign-accented condition as in the native-produced condition. Therefore, no clear conclusion can be drawn from this result whether foreign accents impact the processing of words and nonwords varying in phonotactic probability.

Experiment 4

A fairly clear picture emerges from Experiments 1-3 that the locus of difficulty in recognizing foreign-accented words mainly lies on the lexical level. However, further study is required to examine the precise mechanism of this accent-induced processing difficulty at the lexical level. How exactly do acoustic-phonetic deviations from foreign accents impact the processing at the lexical level? There are several possibilities. First, with greater mismatches between the accented speech input and the native phonological representations in the listeners, the lexical candidates, including both the target word and other competing candidates, might not be as activated as in native speech. Thus, it might take a longer time for activation to pass the threshold for recognition. This possibility does not predict any specific pattern for accuracy rate in recognizing foreign accented speech.

Second, given that the nonnative sound categories are acoustically more variable and distributed in much greater proximity (Sidaras et al., 2009; Wade et al., 2007), words that are not that similar to the target words might also be activated as lexical candidates. Since more words are activated as competing candidates in this case, it might take a longer time for the best match for the accented input to be resolved from the competition. This scenario also predicts that words that are less similar to the target words might also be mistaken as the target words, and the misperceptions from the accented condition might show a greater variability than those from the native condition.

Third, with the foreign accent, the accented pronunciation of the target word might just sound like one of its similar sounding words. In this scenario, that similar sounding word, which seems to be the best matched, might get most highly activated and win out the competition instead of the target word. Thus, this scenario predicts that mostly similar sounding words of the target words would be mistaken as the target words. Also, based on the assumption that only few of the similar sounding words would be consistently best-matched with the accented target word, misperceptions from the accented condition are predicted to show less variability than those from the native condition.

These three scenarios are possible consequences directly resulted from the acoustic-phonetic deviations in foreign accents. There are other possible scenarios that are not directly resulted from accent-induced acoustic-phonetic deviations, but might happen in both native and accented conditions. First, it is possible that listeners respond with a bias or guess strategy towards the most frequently occurring neighbor instead of a neighbor that sounds like the accented pronunciation of the target word. This scenario predicts that the most frequently occurring neighbor of the target word would be mistaken as the target words. Second, it is also possible that words with fewer neighbors than the target words would be more likely to be activated and mistaken as the target words. This hypothesis is based on numerous studies on native-spoken word recognition showing that sparse words are generally retrieved more quickly and accurately than dense words due to fewer competitors during lexical retrieval (Luce & Pisoni, 1998).

To gain insight into the underlying mechanism involved in foreign-accented word recognition, it is important to investigate the set of lexical candidates being activated during the lexical competition stage. Therefore, a perceptual identification task was conducted in Experiment 4 to examine the misperceptions of native and foreign-accented words. In this task, participants were presented with a stimulus word, either native-produced or foreign-accented (against a background of white noise only for the native-produced condition), and were asked to identify it. The perception errors (misperceptions) were words that were most highly activated during the lexical competition and misrecognized as the target words.

One possible way to distinguish the second and third possible mechanisms behind the difficulty in recognizing foreign-accented words is to compare the phonological similarity between the target word and the set of lexical candidates activated during lexical competition. A low similarity between them would imply that the less similar sounding words are the competitors, supporting scenario 2—word candidates that are not closely similar to the target words were mis-activated. A high similarity between the target word and the set of lexical candidates activated would imply that the similar sounding words were the strong competitors, supporting scenario 3— similar sounding words were more highly activated than the target words.

An additional way to distinguish the possible mechanisms behind the difficulty in recognizing foreign-accented words is to compare the variability of the misperceptions from the native and the accented conditions. It would allow us to check whether a variety of words are activated as in scenario 2, or consistently only few similar sounding words that best matched with the accented stimuli was activated as in scenario 3. To test the hypothesis of a word-frequency bias, the percentage of misperceptions that were the most frequently occurring neighbor of the target words was calculated. To test whether words with fewer neighbors than the target words would be more likely to be mistaken as the target words, the mean neighborhood density of all the misperceived target words and all the misperceptions were compared.

Apart from the misperceptions, the transcription accuracy rates for the dense and sparse words across the two accent type conditions were also analyzed. This allowed us to check whether the accent-induced processing costs reflected as longer reaction times in Experiment 1 was also reflected in the accuracy rates in the current experiment. Results mirroring those of Experiment 1 were expected—the recognition of dense words was expected to be more difficult than sparse words; this neighborhood density effect was expected to be strong for the foreign-accented condition relative to the native condition.

Method

Stimuli and design

The same 64 word stimuli from Experiment 1 were used in the present experiment. In an attempt to avoid both floor and ceiling effects on the word recognition scores, white noise of the same signal to noise ratio (S/N) was intended to mix with the stimulus sound files from both the native and foreign-accented conditions. After a series of informal screening and piloting experiments using a quiet listening condition and noisy listening conditions with a variety of S/N ratios, it was found that the piloting listeners experienced substantial difficulties (with an average recognition accuracy rate of 50-60%) in identifying the foreign-accented words even in a quiet listening condition. However, the native-produced words were identified almost perfectly (with an average recognition accuracy rate of 93-98%) in a quiet listening condition. Therefore, to minimize floor effect for the foreign-accented condition and the ceiling effect for the native condition, the author decided to adopt a quiet listening condition and a noisy listening condition for the foreign-accented and native conditions respectively.

White noise with the same duration and relative amplitude as the sound file was added to the 128 stimulus sound files from the two NSs of English only (64 stimulus words x 2 speakers) using the GSU Praat tools (Owren, 2008) in Praat. It yielded a +20dB S/N so that the mean amplitude of the resulting sound files were 20 dB more than that of the white noise, and it resulted in a reasonable range of 80-90% word recognition accuracy for the native condition.

Counterbalancing Procedure

The same counterbalancing procedure used in Experiment 2 was followed in the present experiment. With *accent type* and *speaker* as between-subjects factors and neighborhood density as a within-subjects factor, participants were randomly assigned to one of the four counterbalanced conditions, in which all the stimuli were produced only by one of the four speakers: NMS, NFS, AMS, or AFS. Thus, a given listener only heard stimuli spoken by one of the four speakers and each of the 32 dense and 32 sparse words once. Items were presented in a different randomized order for each participant. Across participants, each stimulus was presented in both native and foreign-accented form and evenly represented by each speaker.

Listeners

Forty native speakers (NS) of American English were recruited from the pool of Introductory Psychology students enrolled at the University of Kansas. The participants received partial credit towards the completion of the course for their participation. All participants were right-handed and reported no history of speech or hearing disorders. None of the participants in the present experiment took part in any of the other experiments in this study.

Procedure

Listeners were tested in a group of up to three persons each time. Each participant was seated in front of an iMac computer in an individual listening station separated by partitions. In the perceptual identification task, each trial begins with the word "READY" appearing on the computer screen for 500 ms. The participants then heard one of the randomly selected stimulus words, either in quite for the foreign-accented condition or embedded in white noise for the native-produced condition, through a set of Beyerdynamic DT 100 headphones at a comfortable listening level. Each stimulus was presented only once. The participants were instructed to use the computer keyboard to enter their response (or their best guess) for each word they heard over the headphones. They were instructed to type "?" if they were absolutely unable to identify the word. The participants were allowed as long as they needed to respond until they finished by hitting the RETURN key, and then the next trial began. Participants were able to see their responses on the computer screen when they were typing and could make corrections to their responses before they hit the RETURN key. The experiment lasted about 10-15 minutes. Prior to the experiment, each participant received five practice trials to become familiar with the task. These practice trials were not included in the data analyses.

Results

Transcription Accuracies

For the perceptual identification task, transcription accuracy rates were the dependent variable of interest. In the data scoring, a response was scored as correct if the phonological transcription of the response and the stimulus was an exact match. Misspelling, transpositions, and typographical errors that involve a single letter in the responses were scored as correct responses in certain conditions: (1) the omission of a letter in a word was scored as a correct response only if the response does not form another English word, (2) the transposition or addition of a single letter in the word was scored as a correct response if the letter is within one key of the target letter on the keyboard. Responses that did not meet the above criteria were scored as incorrect.

To check for any speaker effect, transcription responses were pooled across stimulus items within each speaker condition, yielding a mean percent correct transcription scores for each of the two levels of neighborhood density for each of the four speakers, including NMS, NFS, AMS, and AFS, presented in Table 8. ANOVAs with speaker as a between-subjects factor and neighborhood density as a within-subjects factor were conducted on the mean percent correct transcription scores for participants listening to the native speakers (NMS and NFS) and the foreign-accented speakers (AMS and AFS) respectively. The ANOVAs revealed no significant main effect of speaker for the native condition (F (1, 18) = 3.61, $p > .05$), nor the foreign-accented condition (F (1, 18) =.55, $p > .05$). There was also no significant interaction between speaker and neighborhood density for the native condition (F (1, 18) < .0001, $p > .05$) and for the foreign-accented condition (F (1, 18) =2.5, $p > .05$). Thus, it was assumed that the influence of the speakers' individual idiosyncrasies on the intelligibility of their word productions was minimal; listeners' responses were pooled across speakers during all the data analyses followed.

With participants as the random variable, responses were pooled across speakers and stimulus items, yielding mean percent correct transcription scores in each of the two neighborhood density conditions for each participant. A 2 (accent type: native vs. foreign-accented) x 2 (neighborhood density: dense vs. spare) mixed-design ANOVA, with accent type as a between-subjects factor and neighborhood density as a within-subjects factor, was conducted on the mean percent correct transcription scores. The ANOVA yielded a significant main effect of accent type, F_1 (1, 38) = 267.92, $p > .0001$. Stimulus words spoken by foreign-accented speakers (M = 54.14%, SE = 1.48) were recognized less accurately than those spoken by native speakers (M = 88.44%, SE = 1.48). There were no significant main effect of neighborhood density (F_1 (1, 38)

= 1.016, *p* = .32), nor interaction between neighborhood density and accent type (F_1 (1, 38) = .79, *p* = .379). Hence, there was no significant difference in mean percent correct transcription scores between dense and sparse words regardless of accent type (see Figure 8, M_{dense} = 89.69%, *sd* = 8.78 vs. M_{sparse} = 87.19, *sd* = 8.29 for the native condition; M_{dense} = 54.22%, *sd* = 6.90 vs. M_{sparse} = 54.06%, *sd* = 7.18 for the accented condition).

With items as the random variable, responses were be pooled across speakers and listeners to yield mean percent correct transcription scores in the native and foreign-accented conditions respectively for each item in the dense and sparse conditions. A 2 x 2 mixed-design ANOVA, with accent type as a within-words factor and neighborhood density as a between-words factor, were conducted on the mean percent correct transcription scores. The ANOVA yielded a significant main effect of accent type, F_2 (1, 62) = 54.60, *p* > .0001. Stimulus words spoken by foreign-accented speakers (*M* = 56.41%, *SE* = 4.17) were recognized less accurately than those spoken by native speakers (*M* = 89.22%, *SE* = 1.93). There were no significant main effect of neighborhood density (F_2 (1, 62) =.13, *p* > .05), nor interaction between neighborhood density and accent type (F_2 (1, 62) = .36, *p* > .5). Hence, there was no significant difference in mean percent correct transcription scores between dense and sparse words regardless of accent type (M_{dense} = 89.69%, *sd* = 15.81 vs. M_{sparse} = 88.75, *sd* = 15.13 for the native condition; M_{dense} = 54.22%, *sd* = 29.41 vs. M_{sparse} = 58.59%, *sd* = 36.90 for the accented condition).

Analysis of Misperceptions

The phonological characteristics of the misperceptions (*i.e.,* incorrect responses) were analyzed and coded according to the degree of phonological similarity with the target word: (1) low similarity defined as one or no phoneme overlap, or (2) high similarity defined as two phoneme overlap. The observed frequency counts of misperceptions were pooled across speakers, listeners and items, and tallied for each of the eight cells in a 2 (accent type: native vs. foreign-accented) x 2 (neighborhood density: dense vs. spare) x 2 (phonological similarity with the target word: low vs. similar) contingency table, see Table 9.

In Experiments 1 to 3, ANOVAs were used to evaluate association and interactive effect of two or more categorical predictor variables (e.g., neighborhood density, phonotactic probability, and accent type) on a continuous outcome variable (e.g., reaction time and accuracy rates), thus testing the difference between the means of two or more groups. However, all three variables in the current experiment— accent type, neighborhood density, and phonological similarity— were categorical in nature. Therefore, log-linear analysis, which is commonly used to evaluate association and interaction patterns among a set of three or more categorical variables, was used to analyze the 2 x 2 x 2 contingency table containing the frequency counts of misperceptions. It modeled how the two qualitative explanatory variables, accent type (AT) and neighborhood density (ND), predicted the degree of phonological similarity (PS) between the misperceptions and the targets. A backward elimination procedure with the likelihood-ratio model comparison method was used to test the main effects and interaction in the log-linear model.

A saturated model with all the main effects and all the interactions, symbolized by AT x ND * PS, listed as model M1 in Table 11, was first fitted. Removing the 3-way interaction term yielded the simpler model M2 with all the two-way interactions (AT x ND, AT x PS, ND x PS) and main effects. The likelihood-ratio statistics comparing the two models (M1 and M2) equaled the difference in deviances $G^2 = .52$, $df = 1$, $p = .47$, suggesting that the three-way interaction term was not significant.

The simplification process continued by successively eliminating the 2-way interaction terms in the model that was the least significant (with the largest p-values). Taking out the interaction between accent type and neighborhood density (M3a) did not significantly decrease the model fit, $G^2 = .15$, $df = 1$, $p = .7$. This suggested that the conditional independence between accent type and neighborhood density, controlling for the degree of phonological similarity between the misperceptions and the targets. That is, there is no association between accent type and neighborhood density when adjusted for the degree of phonological similarity between the misperceptions and the targets. The elimination process proceeded as showed in Table 11. After the whole elimination procedure, the final model was M3a that it contained all the main effects, the accent type x phonological similarity interaction, and the neighborhood density x phonological similarity interaction. Therefore, it suggested significant accent type -

phonological similarity and neighborhood density -phonological similarity partial associations.

Based on the final fitted model, the estimated accent type -phonological similarity conditional log odds ratio equaled .542, *SE* = .177. A 95% confidence interval for the true conditional log odds ratio was (.196, .888), yielding (1.22, 2.43) for the true conditional odds ratio. This suggested a significant positive association between accent type and phonological similarity, controlling for neighborhood density. That meant, for both dense and sparse neighborhood density conditions, the odds for foreign-accented words to have highly similar competitors was 1.72 (i.e., $e^{.542}$) times the odds for native-produced words to have highly similar competitors.

The estimated neighborhood density -phonological similarity conditional log odds ratio equaled 1.245, *SE* = .055. A 95% confidence interval for the true conditional log odds ratio was (1.136, 1.353), yielding (3.11, 3.87) for the true conditional odds ratio. This suggested a significant positive association between neighborhood density and phonological similarity, controlling for accent type. That meant, for both native and foreign-accented conditions, the odds for dense words to have highly similar competitors was 3.47 (i.e., $e^{1.245}$) times the odds for sparse words to have highly similar sounding competitors.

The consistency of the misperceptions was also analyzed to check whether a variety of words were activated, or consistently only few similar sounding words that best matched with the accented stimuli were activated. A misperception was coded as consistent if there were at least one or more misperceptions for the same target word that shared its phonology. Otherwise, the misperception was coded as inconsistent. The observed frequency counts of misperceptions were pooled across speakers, listeners and items, and tallied for each of the four cells in a 2 (accent type: native vs. foreign-accented) x2 (consistency: consistent vs. not consistent) contingency table, see Table 10. A Chi-square test was used to analyze the 2 x 2 contingency table containing the frequency counts of misperceptions. It seemed that misperceptions from the accented condition were more likely to be consistent than were misperceptions from the native condition, χ^2 (1, N = 720) = 46.39, p < .0001. In other words, listeners were more likely to come up with the same misperception for the same target word in the accented condition than in the native condition.

To test the hypothesis of a word-frequency bias, the percentage of misperceptions that was the most frequently occurring neighbor of the target words was calculated. Only 8.97% and 3.13% of the misperceptions were the most frequently occurring neighbor of the target words for the native and the accented condition respectively. Therefore, it seemed that listeners did not respond with a bias towards the most frequently occurring neighbor.

To test whether words with fewer neighbors than the target words would be more likely to be mistaken as the target words, the mean neighborhood density of all the misperceived target words and all the misperceptions were compared. Stimulus items that induced no perception errors in the participants were not included in this analysis. The mean neighborhood density of all the (misperceived) stimuli and all the corresponding misperceptions were subjected to a one-way ANOVA. For the native condition, no significant difference was found between the means neighborhood density of all the misperceived stimuli ($M = 22.34$, $sd = 5.80$) and all the corresponding misperceptions ($M = 20.89$, $sd = 7.41$, $F(1, 74) = .907$, $p = .344$). For the accented condition, no significant difference was found between the means neighborhood density of all the misperceived stimuli ($M = 22.40$, $sd = 6.12$) and all the corresponding misperceptions ($M = 21.97$, $sd = 6.57$, $F(1, 114) = .13$, $p = .78$).

Discussion

A reliable main effect of accent type was found on the transcription accuracies. Even though the native stimuli were presented in a noisy listening condition, whereas the foreign-accented stimuli were presented in a quiet listening condition, the native stimuli ($M = 88.44\%$) yielded significantly higher intelligibility scores than the accented stimuli ($M = 54.14\%$). Compared to the masking effect induced by noise on the native stimuli, the foreign accents induced a significantly greater word recognition difficulty in the native listeners. This result confirmed that our accented speakers had a markedly strong foreign accent. Therefore, it supported the effectiveness of the accent-type manipulation.

Inconsistent with the findings from the lexical decision task in Experiment 1, results from the current experiment showed no significant main effect of neighborhood density, or two-way interaction between neighborhood density and accent type. The current result failed to show any significant difference in transcription accuracy between dense and sparse words in either the native, or the foreign-accented conditions. Like Imai et al's study (2005), the current result failed to replicate previous findings of a reliable neighborhood density effect consistently found in a perceptual identification task (i.e., an intelligibility task with noise-degraded stimuli) with native-produced words and native listeners (Luce & Pisoni, 1998; Vitevitch, Stamer, & Sereno, 2008). However, the current experiment also failed to replicate the significant neighborhood density effect found during foreign-accented word recognition in Imai et al's study (2005). Remember from Imai et al's study (2005) that a significant neighborhood density effect was found during foreign-accented word recognition in native listeners, but not during native word recognition. The current results also failed to show such a differential influence of neighborhood density on the intelligibility of the native-produced and foreign-accented words.

There might be many reasons for obtaining these null results. One possible reason for the lack of an interaction between neighborhood density and accent type is that the foreign accent was too strong in this study that it resulted in a floor effect. The foreign-accented stimuli were presented in noise (S/N ratios of +14 dB) in Imai et al's study (2005), in which a significant neighborhood density effect was found with around 60% and 40% recognition accuracies for sparse and dense words respectively. However, in the current experiment, the foreign-accented stimuli were presented without noise, and only around 54% transcription accuracies were obtained in both sparse and dense conditions. Comparing the presentation conditions and transcription accuracies across the two studies, it seems that the foreign accents incurred a relatively weaker adverse effect on word recognition in Imai et al's study than in the current study. This suggested that stronger foreign accents were used in the current study, which might induce a greater extent of pronunciation deviations, making word recognition markedly difficult for the native listeners regardless of neighborhood density. Another possibility is that participants might respond to the stimuli using some sort of sophisticated guessing strategy as they were given as much time as they needed to recognize the words, especially that the foreign accents used in the current study were strong.

Misperception analysis was conducted to study the activated alternative word candidates during lexical competition. There was a significant positive association between accent type and phonological similarity, controlling for neighborhood density. That meant for both dense and sparse neighborhood density conditions, the probability for the misperceptions to share two phonemes with the target words was higher in the foreign-accented condition than in the native-produced condition. This suggested that during foreign-accented word recognition, the larger mismatches between listeners' phonological representations and the speech input do not tend to activate competing words that are less similar to the target words. Instead, the mis-activated alternative word candidates during foreign-accented word recognition tended to more closely resemble the target words than those during native word recognition.

Analysis of the consistency of the misperceptions showed that misperceptions from the accented condition were more likely to be consistent than were misperceptions from the native condition. This suggested that listeners were more likely to come up with the same misperception for the same target word in the accented condition than in the native condition. This implies that only few similar sounding words that best matched with the accented stimuli were consistently activated in the presence of foreign accents. This ruled out the possibility that a wide variety of words were activated by mismatches from foreign accents.

Taken together, this is more consistent with the third hypothesis mentioned earlier that the target words are facing increased competition from some of its similar sounding words at the stage of lexical competition during foreign-accented word recognition. If similar sounding words are mis-activated during lexical retrieval, these competitors might prolong the processing time necessary to identify the target word in several possible ways, depending on the competition mechanism in the spoken word recognition models. It is possible that these similar sounding words might suppress the target word under its resting level of activation, resulting in delayed activation of the target word. Due to their high phonological similarity to the target word, these similar sounding words might also be reluctantly suppressed by lateral inhibition from other competitors, thus posing increased competition for the target word (McClelland & Elman, 1986; Norris, 1994).

Increased activation of multiple similar sounding candidates might also lower the probability of identification of the target word indirectly at the stage of decision.

For example, in the neighborhood activation model (NAM), the probability of identification of the target word depends on the activation level of the target word, the combined level of activation of its neighbors, and the frequency information (Luce & Pisoni, 1998). As a result, mis-activation of similar sounding words poses increased lexical competition for the target word. This possibly contributed to the greater processing costs for foreign-accented word recognition that mainly originated at the lexical level of processing.

There was also a significant positive association between neighborhood density and phonological similarity, controlling for accent type. That meant, for both native and foreign-accented conditions, the probability that the misperceptions have two phonemes overlap with the target words was higher for dense words than for sparse words. Consider that there are a relative greater number of similar sounding lexical words available to act as alternative candidates in the case of dense words than sparse words. With many similar sounding lexical words available as alternative candidates for word recovery, there is a higher chance for the misperceptions of dense target word to be one of its phonological neighbors. However, for sparse words, as there are not as many similar sounding lexical words available as alternative candidates, the words that are mis-activated may not be phonologically similar to the target words.

Regarding the hypothesis of a word-frequency bias, a markedly small percentage of misperceptions were the most frequently occurring neighbor of the target word. Therefore, it is unlikely that listeners responded with a bias towards the most frequently occurring neighbor instead of a neighbor that sounds like the accented pronunciation of the target word. To test whether words with fewer neighbors than the target words would be more likely to be mistaken as the target words, the mean neighborhood density of all the misperceived target words and all the misperceptions were compared. Regardless of accent type, no significant difference was found between the mean neighborhood density of all the misperceived target words and all the corresponding misperceptions. This implies that words with fewer neighbors than the target words have no advantage to be activated and mistaken as the target words.

In sum, the transcription accuracy results failed to replicate the differential influence of neighborhood density on the intelligibility of the foreign-accented words shown in Imai et al's study (2005) or in Experiment 1 (lexical decision task). However, the findings from the misperception analysis gave us insights into the set of lexical candidates being activated during the lexical competition stage. Compared to sparse words, the competing candidates for dense words tend to be more phonologically similar to the target words. Compared to native-produced words, the competing candidates for foreign-accented words tend to be more consistent and more phonologically similar to the target words. This supports the hypothesis that similar sounding words of the target words might get more highly activated in the presence of foreign accents and win out the competition instead of the target word, contributing to a greater foreign-accented word recognition difficulty originating at the lexical level.

General Discussion

Implications for Foreign-accented Word Recognition in Human

This series of experiments utilize isolated word and nonword stimuli to directly examine the effects of neighborhood density and phonotactic probability on foreign-accented word recognition. The designs of all four experiments are summarized in Table 12. Findings from these experiments shed light on several signature problems in the field of foreign-accented speech research: (1) would certain foreign-accented words be more difficulty to recognize depending on their sub-lexical and lexical characteristics? (2) where is the locus of difficulty in recognizing foreign-accented words? (3) how do accent-induced mismatches impact the lexical retrieval process specifically?

In the current study, experiments that emphasized the lexical level of processing (Experiments 1 & 4) consistently showed increased reaction times and lower accuracy rates to the foreign-accented rather than native-produced stimuli, demonstrating reduced word recognition performance due to foreign accents. These results are consistent with previous studies, which demonstrated extra processing cost for foreign-accented speech in terms of lower intelligibility and longer processing time (Munro & Derwing, 1995a, 1995b).

However, in experiments that emphasized sub-lexical level of processing (Experiments 2 & 3), no extra processing costs, in either reaction times or accuracy rates, were found for the foreign-accented stimuli. These results suggest that the sub-lexical processing of the foreign-accented words is as good as that of the native-produced words. This result might seem surprising given previous evidence of foreign accents disrupting sub-lexical processing, such as a higher confusability on vowel recognition in foreign-accented speech than native speech (Sidaras et al., 2009; Wade et al., 2007). However, a previous study by Burki-Cohen et al. (2001) mentioned earlier also didn't find any significant processing costs for foreign-accented stimuli relative to native stimuli in a task that emphasizes sub-lexical processing—the phoneme monitoring task— under ideal listening condition. Therefore, more research is needed to study how foreign accents impact the sub-lexical level of processing, and whether the effect is task-specific or dependent on the listening condition. Taken together, the present results suggest that foreign accents primarily impair the lexical level of processing, and have little to no influence on the sub-lexical level of processing.

Previous studies by Burki-Cohen et al (2001) and Imai et al (2005) have showed the influences of lexical factors, including word frequency and neighborhood density, on foreign-accented word recognition, but only under adverse listening conditions. In this study, using non-degraded stimuli, Experiment 1 found increased accent-induced processing costs, especially for words with many similar sounding words, implying that foreign-accented word recognition is more undermined for words from dense neighborhoods. Using a reaction time measure, this experiment replicated the enhanced neighborhood density effect found in the foreign-accented condition in Imai's et al (2005) study, which used transcription accuracy measures. Most importantly, the current study further demonstrated that the lexical effect was extended to optimal listening conditions. This result suggests that native listeners increase their reliance on top-down lexical knowledge during foreign-accented word recognition. In this case, recognition of dense words would be more difficult than sparse words as the similar sounding words of dense target words would be more activated leading to increased competition for recognition.

Regarding the sub-lexical factors on foreign-accented word recognition, Experiment 3 failed to demonstrate any significant difference in the sub-lexical processing of foreign-accented words and nonwords with high and low phonotactic probability. Given that the phonotactic probability effect was also insignificant in the native condition, it is not clear whether the effect in the foreign-accented condition was not strong enough to be detected, or it manifested the same way as in the native condition. Thus, no clear conclusion can be drawn on whether native listeners take advantage of the phonotactic information during sub-lexical processing of foreign-accented stimuli, or whether they experience greater difficulty in sub-lexical processing of accented items with less common segments and sequences of segments.

Taken together, these results imply that native listeners tend to increase their use of lexical knowledge to compensate the sub-optimal accented speech input for word recognition. Munro and Derwing (1995b) speculated that top-down information, such as lexical-constraining context, might be used to restore particular phonemes, segments, words or phrases that were misunderstood. Results from the current study suggest that knowledge of lexical word form is used to resolve the mismatches between accented speech input and native phonological representations.

Of particular interest in this study is locating the locus of the difficulty associated with foreign-accented word recognition by comparing the pattern of results across Experiments 1 and 2, which emphasized the lexical and sub-lexical levels of processing respectively. Results from Experiment 1 showed an overall processing delay for foreign-accented words, as well as an increased processing delay for accented dense words relative to sparse words, in the LD task. However, there is not any overall processing delay observed for foreign-accented stimuli, nor increased processing delay particularly for accented dense words in the AX task. Therefore, the accent-related processing costs observed in the LD task in Experiment 1 were not evident in the AX task, which emphasized sub-lexical processing. This suggests that foreign accents do not seem to impact the sub-lexical level of processing. The overall recognition difficulty of foreign-accented stimuli observed in Experiment 1 mainly stems from the lexical level. The differentially increased processing difficulty for accented dense words in the lexical decision task also seemed to mainly originate from the lexical level.

Another unique contribution of the current study is to shed light on how accent-induced mismatches impact the lexical retrieval process by studying the set of lexical candidates being activated during the lexical competition stage. The misperception analysis from Experiment 4 found a significant positive association between accent type and phonological similarity, controlling for neighborhood density. That meant there was a higher probability for the misperceptions to be highly similar to the target words in the foreign-accented condition than in the native-produced condition, regardless of neighborhood density. This suggested that the larger mismatches between listeners' phonological representations and the foreign-accented speech input do not tend to activate more competing words, which are less similar to the target words.

Instead, the alternative word candidates activated during foreign-accented word recognition are still restricted to those words that are highly similar to the target words. However, these similar sounding words might be more activated than normal by the mismatches induced by foreign accents compared to native-produced word recognition. As a result, longer time is needed to resolve the increased competition at the lexical level, and there is a higher chance for similar sounding words to be mistaken as the target words during foreign-accented word recognition. This is more consistent with the hypothesis that the target words are facing increased competition from some of its similar sounding words at the stage of lexical competition during foreign-accented word recognition. This possibly contributed to the greater processing costs for foreign-accented word recognition that mainly originated at the lexical level of processing.

To summarize, the experiments fill in large holes in our understanding of foreign-accented word recognition, especially regarding how recognition difficulties arise from the lexical level of processing, and how listeners use top-down lexical word form knowledge to compensate for mismatches induced by accented speech inputs. The present results also increase our fundamental knowledge about how native listeners compensate for the variability associated with accented speech, and thus provide insight into the perceptual mechanisms that underlie foreign-accented word recognition.

Implications for Spoken Word Recognition Models

The signature problem in speech perception is how listeners achieve perceptual constancy given the great variability in the acoustic realization of lexical items introduced by speakers and speech contexts. Foreign-accented speech deviates from native speech in subsegmental, segmental, and suprasegmental levels (Flege, 1984) and contains greater within- and across-speaker variability (Wade et al., 2007). Thus, it is important to evaluate the validity of current spoken word recognition models in accounting for the behavioral data of foreign-accented word recognition observed in the current study. If these models are able to correctly predict/account for the observed word recognition behavior, this will strengthen the theory and the underlying assumptions of the models. If these models cannot correctly predict the observed behavior, this suggests that these models might fail to incorporate important mechanisms to deal with these accent-related acoustic deviations and variability.

Spoken word recognition involves mapping of the incoming acoustic speech signals onto the representations of words stored in our memory so that the utterances can be interpreted. There are two basic views on the nature of the lexical representations— the abstract and the exemplar-based— to deal with this speech variability effectively. The exemplar-based view assumes a collection of detailed acoustic traces representing individual words in the lexicon (Goldinger, 1998). Lexical access then involves matching the incoming speech input to collections of acoustic traces stored in the mental lexicon.

On the other hand, the abstractionist view assumes the lexical entries to be a set of abstract, ideal, and modality-free representations. Therefore, the perception system is assumed to filter and discard the surface details tangential to the word identity through a process known as normalization, resulting in only canonical mental representations at the prelexical level for subsequent lexical processing. Therefore, the speech recognition process in abstractionist theories involves mapping incoming acoustic signal onto abstract prelexical representations, which are then mapped onto the lexical representation stored in the form of a sequence of prelexical units. The abstractionist view is prominent in many of the current models of spoken word recognition, including TRACE (McClelland & Elman, 1986), Shortlist (Norris, 1994), and PARSYN (a connectionist implementation of the NAM with an addition of an explicit sub-lexical processing level; Luce, Goldinger, Auer, & Vitevitch, 2000). These influential models of human spoken word recognition have been implemented as computational models, which aim to simulate and explain empirical data related to the human speech recognition process.

Adopting the abstractionist view on lexical representations, current spoken word recognition models, TRACE (McClelland & Elman, 1986), Shortlist (Norris, 1994), and PARSYN (Luce et al., 2000) all assume speech input to be perfect and free of acoustic variability of any kind. The computational implementations of these models all fail to specify how the acoustic signal is converted into abstract prelexical representations. Instead of taking an acoustic signal (i.e., real speech signal) as input, these models all take in some kind of prelexical representations as input, which are either artificial or manually pre-transcribed by human (Scharenborg, 2007). Speech input is assumed to be a sequence of discrete phonemes in Shortlist (Norris, 1994), a form of pseudospectral representations based on acoustic-phonetic features in TRACE (McClelland & Elman, 1986), and position-specific allophones in PARSYN (Luce et al., 2000). Hence, it is not clear whether the acoustic-phonetic alternations in foreign-accented speech input are being maintained, and how the foreign-accented acoustic signal is converted into the form of prelexical input specified by each of these models. The lack of clarity in how the foreign-accented acoustic signal is converted into the form of prelexical input specified by each of these models means that the predictions discussed below are speculative at best.

Although the specific form of prelexical representation and input are different, current spoken word recognition models, TRACE (McClelland & Elman, 1986), Shortlist (Norris, 1994), and PARSYN (Luce et al., 2000), all posit prelexical and lexical representation layers with a two-stage processing—an activation stage and a decision stage based on lexical competition. What really distinguishes these models is their assumptions on the flow of activation between the prelexical and lexical levels. Specifically, during foreign-accented word recognition, it is how the higher-level lexical influence is implemented to recover the word identity from the effects of foreign accents. This implies that the particularly crucial higher-level lexical influence in foreign-accented word recognition is implemented differently in these models. More importantly, the different mechanisms involved in these models imply different predictions on the recognition performances for foreign-accented speech, such as the locus of processing difficulty and the types of recognition errors.

Imagine a scenario that a strong foreign accent is present, and the mismatch between the speech signals and pre-lexical representations is so severe to the extent that one or more phonemes are indistinguishable at the phoneme level alone. Then top-down lexical information is needed to "resolve" the disrupted pre-lexical processing and the identity of the target word. Lexical influence in interactive TRACE is mediated by bi-directional flow of information, whereas lexical influence in both Shortlist and NAM is restricted to post-perceptual decision processes. More specifically, in TRACE, top-down biasing activation from the lexical level is allowed and it boosts the activation for the lexically consistent phonemes at the pre-lexical level, replacing the insufficient pre-lexical representations from speech input and resulting in bias towards lexically favored phoneme decisions. That is, TRACE allows the feedback of lexical knowledge to alter pre-lexical processing.

In contrast, top-down feedback of lexical knowledge is not allowed in Shortlist and is absent in PARSYN to alter pre-lexical processing (Luce et al., 2000; Norris, 1994); consequently, the pre-lexical representations of Shortlist and PARSYN stay ambiguous, as is from input, and are intact from top-down lexical influence. Word recognition in Shortlist and PARSYN is derived from the competition among the lexical candidates that were activated based on bottom-up input. Therefore, the accent-induced sub-lexical ambiguity can be “resolved” at either the sub-lexical or lexical levels in TRACE, but only at the lexical level in Shortlist and PARSYN. For TRACE, the interactive flow of activation between the prelexical and lexical levels that is responsible for restoring the accent-induced acoustic-phonetic alternations probably might account for the longer processing time in accented word recognition. For Shortlist and PARSYN, the longer processing time for accented word recognition may be accounted for by the increased difficulty in the lexical competition process on the basis of insufficient pre-lexical information.

These predictions imply that the major locus of foreign-accented word recognition difficulty probably lies at the lexical level according to Shortlist and PARSYN, but it could be at either the pre-lexical or lexical level or both according to TRACE. The results from the current study indicate that the major locus of foreign-accented word recognition difficulty lie at the lexical level. Thus, it is consistent with the predictions from Shortlist, PARSYN, and TRACE. However, without examining an actual simulation, the above predictions about how the complex computational models might perform with foreign-accented speech should be taken with caution (Lewandowsky, 1993).

Moreover, the different constraints the models place on the information flow between representation levels also has implication for the lexical candidates that are activated during lexical competition. Shortlist activates a set of lexical candidates based on only the bottom-up input, whereas TRACE tends to activate a set of lexical candidates due to the interactive activation between the phoneme and word levels. As a result, when strong foreign accents are present so that the speech signals are highly ambiguous, the top-down interactive influence in TRACE will be enhanced to resolve the ambiguity. In this case, the ambiguous representation of the accented input at the sub-lexical level might tend to be replaced by the stronger top-down feedback from the lexical level. Thus, the only representation of the accented input that is available in the model will be lost. This may lead to a set of lexical candidates, which are mainly driven by lexical knowledge instead of the speech input, to be activated and increase the risk of "hallucinating" in TRACE (Norris, McQueen, & Cutler, 2000). Losing representation of the accented input and the "hallucination" in TRACE might explain why listeners are reluctant to give-up on the activated lexical candidates and retrieve something else instead. In contrast, the lexical candidates in Shortlist and PARSYN are always input-driven.

Hence, in the presence of a strong foreign accent, these models predict different types of recognition errors. TRACE tends to predict errors that may not be phonologically similar to the intended target words; however, both Shortlist and PARSYN tend to predict errors that are highly phonologically similar to the intended target words. Results from the misperception analysis in Experiment 4 showed a high degree of phonological similarity between the misperceptions and the target words in the foreign-accented condition. Therefore, both Shortlist and PARSYN seem to provide a better account for foreign-accented speech recognition errors than TRACE.

In sum, the influential models of spoken word recognition, including Shortlist, PARSYN, and TRACE, all assume the abstractionist view on lexical representations that involves mapping incoming acoustic signal onto abstract prelexical representations, which are then mapped onto the lexical representation stored in the lexicon. Their computational models all do not specify the mechanism in converting the acoustic signal into the form of symbolic prelexical representation that is taken as input. Hence, it is not clear how the accent-related acoustic-phonetic alternations in input are being converted into prelexical input specified by these models. Moreover, these three models make different assumptions on whether there is top-down feedback from lexical knowledge to the prelexical lexical. Consistent to the findings from the current study, Shortlist and PARSYN predict the major locus of foreign-accented word recognition difficulty to be at the lexical level, as well as a high phonological similarity between the misperceptions and the target words. In contrast to the current results, TRACE predicts misperceptions to be less phonologically similar to the intended target words. Therefore, Shortlist and PARSYN seems to provide a better account for foreign-accented word recognition observed in this study.

Implications for Foreign-accented Word Recognition in Machine

Foreign-accented words are difficult for humans to recognize, but recognition performance is even worse in automatic speech recognition (ASR) systems. For example, the average word recognition accuracy for SpeechRater, which is a state-of-the-art automatic scoring system for non-native spontaneous speech from the practice tests of the Test of English as a Foreign Language (TOEFL), was only around 50%

(Zechner, Higgins, Xi, & Williamson, 2009). Recognition performance in ASR systems were found to degrade more for foreign-accented speech compared to native speech (Lawson, Harris, & Grieco, 2003).

The challenges that human listeners and ASR face in recognizing foreign-accented words seem to be similar. Foreign-accented speech is less homogenous than native speech with greater within- and across-speaker variability (Wade et al., 2007). The acoustic alternations intrinsic to the foreign-accented speech signal depend on the native language of the accented speaker, as well as the level of his proficiency. This increased variability due to foreign accents is difficult to handle in both human and ASR systems.

An ASR system typically consists of four steps— feature extraction, acoustic modeling, language model, and word searching (Scharenborg, 2007). One ASR approach to model foreign accents is to use multiple acoustic models, a combination of phone models of both the native and target languages to encode non-native phonemes (Bartkova & Jouvet, 2007; Benzeghiba et al., 2007). Another way is to introduce pronunciation variants to encode one or more possible pronunciation variants for each recognition unit (Bartkova & Jouvet, 2007; Benzeghiba et al., 2007). It is worth noting that ASR researchers and engineers aim to use statistical models to develop algorithms that support an efficient and robust performance of the ASR system, regardless of its psychological plausibility. Hence, the representations and algorithms involved in ASR may not be psychologically plausible, and are implemented very differently from human processors, including the way that ASR handles acoustic alternation induced by foreign accents.

Given that ASR has not been even approaching human-like performance regardless of its substantial improvement in the last decades, there is an increasing interests for engineers to improve ASR performance by incorporating the essential knowledge from psycholinguistic research in HSR (human speech recognition) (Moore, 2007; Scharenborg, 2007; ten Bosch & Kirchhoff, 2007). Actually, the decoding processes in HSR and ASR are highly similar in terms of their functions. For example, both the humans and machines have to convert audio input into some forms of abstract representation, use a mapping mechanism to match the input signal to stored representations of words, and use competition to evaluate the best matched word (ten Bosch & Kirchhoff, 2007). Therefore, researchers from the fields of HSR and ASR started to collaborate and exchange knowledge regardless that both fields have markedly different goals (Scharenborg, 2007; Scharenborg, Norris, ten Bosch, & McQueen, 2005)

Research has been done to compare human-machine performance to investigate what leads to human listeners' superior recognition performance over ASR systems, in the hope of improving the ASR system. For example, recognition performance, in terms of accuracy rates, of human listeners and machines has been compared by Lippmann (1997) using several different speech corpora ranging from isolated words to spontaneous conversation. Generally, machine error rates are significantly higher than human error rates. The difference in human-machine performance further increases in noise and in spontaneous speech, suggesting that human listeners use higher-level knowledge to help with challenging speech recognition task. In the absence of high-level grammatical information, human listeners still maintain superior performance over machines in recognizing nonsense syllables and sentences, suggesting that the human-machine performance gap is due to superiority of feature extraction and representations in human listeners (Lippmann, 1997).

Similarly, the speech corpus of foreign-accented and native-produced words used in the current study could be subjected to ASR for evaluating the human-machine performance gap in foreign-accented word recognition. Comparing the human-machine performance, in terms of accuracy rates, could give us insight into how flexible human listeners and ASR systems are to variability induced by foreign accents. More importantly, comparing the pattern of recognition errors, such as types of substitutions, insertions, deletions present in the recognition errors, made by human and machine could tell us how the error recovery process might operate differently in human listeners and ASR systems. This might also give us insight into which properties of the error recovery process in HSR is relevant for the improvement of ASR systems. Moreover, the psycholinguistic data from the current study gave us insights into how human listeners process foreign-accented words, especially regarding how higher-level lexical knowledge might be used to recover the identity of the accented words. This knowledge about HSR might help direct ASR engineers in incorporating a similar top-down error recovery principle in the ASR systems for addressing speech variability due to foreign accents.

Sociolinguistic Factors affecting Foreign-accented Speech Perception

The current study adopted the psycholinguistic approach to investigate foreign-accented speech processing, which exclusively focuses on influences from the lower-level acoustic-phonological deviations. On the other hand, the sociolinguistic approach focuses on influences from the higher-level social cognition. Sociolinguistic researchers have studied how listeners' stereotypes and social biases against foreign-accented speakers could dramatically influence the listeners' perception and evaluation of the accented speech.

Extensive sociolinguistic research has been done on language attitudes, exploring the relationship between foreign accents and listeners' stereotypes about accented speakers. Degree of foreign accent was found to be strongly related to listeners' evaluation of the accented speakers, demonstrating that foreign accents do play a prominent role in triggering the negative stereotypes (Brennan & Brennan, 1981; Ryan, Carranza, & Moffie, 1977). For example, Ryan and colleagues (1977) presented native speakers of American English with different Spanish-accented English audio readings representing a wide range of accentedness and asked for ratings on the speakers' attributes, including status (eventual occupation), solidarity (friendship), and the speech characteristics, including accentedness, pleasantness, and fluency. It was found that the stronger the foreign accents, the more negatively the ratings were on the speech characteristics and the speakers’ attributes. The listeners evaluated the speakers differently in response to different levels of foreign accent, implying that foreign accents are crucial in prompting listeners’ stereotypes about the speakers.

A deep understanding of the cognitive mechanisms involved in foreign-accented speech perception can only be achieved by integrating the psycholinguistic and sociolinguistic perspectives together— in the context of lower-level speech perception and higher-level social cognition. Stereotyping of foreign-accented speakers is an important point in the context of foreign-accented speech perception, given that this higher-level social cognition is simultaneously triggered in the native listeners when they perceive foreign accents in the speech signals, and it may interact in complex ways with the listeners’ speech processing system. As a consequence, further research is required to explore the extent to which the recognition difficulty arises from the social-cognitive and speech processing mechanism respectively.

One approach to this issue is to test the influence of different foreign accents on word recognition. If foreign accents do play an important role in triggering listeners’ stereotypes, and different foreign accents probably trigger different stereotypes specific to the speakers’ original language and nation groups, then the mediating effects of different foreign accents on accented speech perception would probably be different. Following this speculation, native listeners’ spoken word recognition performance is predicted to be differentially influenced by different foreign accents, depending on the different stereotypes being triggered by the different foreign accents.

Previous studies have explored whether different stereotypes are triggered by different foreign accents. Delamere (1996) investigated how American native English speakers evaluate the personality traits of speakers with different foreign accents, including Farsi, Spanish, Malay, Arabic, and French. In that study, a matched-guise technique was adopted so that the text was read by the same speaker twice, one with grammatical errors and one without, which were then presented to the listeners as speech from different speakers. When there were errors in the non-native speech, listeners' ratings of the five speakers with different foreign accents were very similar and differentiated on three traits only. When the nonnative speech contained no errors, listeners tended to rate the five speakers differentially on all the 15 traits. It implied that speech errors might be a salient marker for triggering the general "foreigner" stereotype, whereas error-free speech permitted other salient features to stand out and triggered more stereotypes related to specific language groups. Given that the grammatical errors in the error condition are also commonly found in foreign-accented speech in real life, it is not clear whether the results in the error-free condition (i.e., different foreign accents triggering different stereotypes about the language groups) can be generalized to real life situations.

Gill (1994) also examined how the native North American accent and different foreign accents, including British and Malaysian, affected native listeners' evaluation of the speakers and their speech. The native listeners assigned the most favorable ratings to the native (American) accent, and gave the most favorable attitude evaluations to the native speaker. More favorable ratings were assigned to the British accent than the Malaysian accent, but there was no difference between the listeners' evaluations of the two foreign speakers on all measured dimensions, including socio-intellectual, dynamism, and aesthetic. Of greater significance, comprehension performance was the best for the native speech, whereas no difference was found between the two foreign-accented speeches. Also, a measure of listeners' predispositions to stereotypes did not show a significant effect on the perception of the foreign speakers, suggesting that stereotyping in listeners did not factor into the speaker evaluations, or affect those attitude dimensions measured in that study.

Taken together the findings from Gill's (1994) and Delamere's (1996) studies, it is still not clear whether different foreign accents trigger a general "foreigners" stereotype or specific stereotypes related to the speakers' nation and language groups. Thus, the prediction from the sociolinguistic perspective is not clear whether different foreign accents interfere with word recognition differently or not. The current study investigated the Spanish accent as it is very common in North America so that the current findings are more generalizable to the real world. Also, concerns have been raised against teachers with a heavy Spanish accent since the Arizona Immigration law was passed in 2010. Given the generally negative stereotypes associated with the Spanish accent, it is interesting to replicate the current study using other foreign accents that are associated with neutral or positive stereotypes.

Shifting to the psycholinguistic perspective, different foreign accents mean different sets of systematic accent-general variations, which arise from the interaction of the phonological structures of the speakers' L1 and L2. Hence, different foreign accents might result in different extent of word recognition disruption and different subsets of words being prone to recognition disruption, depending on the differences in the phonetic inventories between the speakers' L1 and L2. Therefore, replicate the current study using other foreign accents will give us a deeper and more coherent understanding of the social-cognitive and language processing mechanisms involved in foreign-accented speech perception.

Conclusions

A proper solution to accent-related communication costs requires a deep understanding of the problems on the side of both the foreign-accented speakers and the native listeners. The current research advanced our knowledge of the origin of accented word recognition difficulties in native listeners, as well as the specific impact of accent-induced mismatches on the lexical retrieval process. The behavioral data from the current study also provided a good testing case for evaluating how well the current models of spoken word recognition in accounting for human foreign-accented word recognition. The speech corpus could also be used for evaluating and comparing recognition performance from human listeners and ASR systems, and provide insights into increasing the tolerance of ASR systems to accent-related acoustic variability.

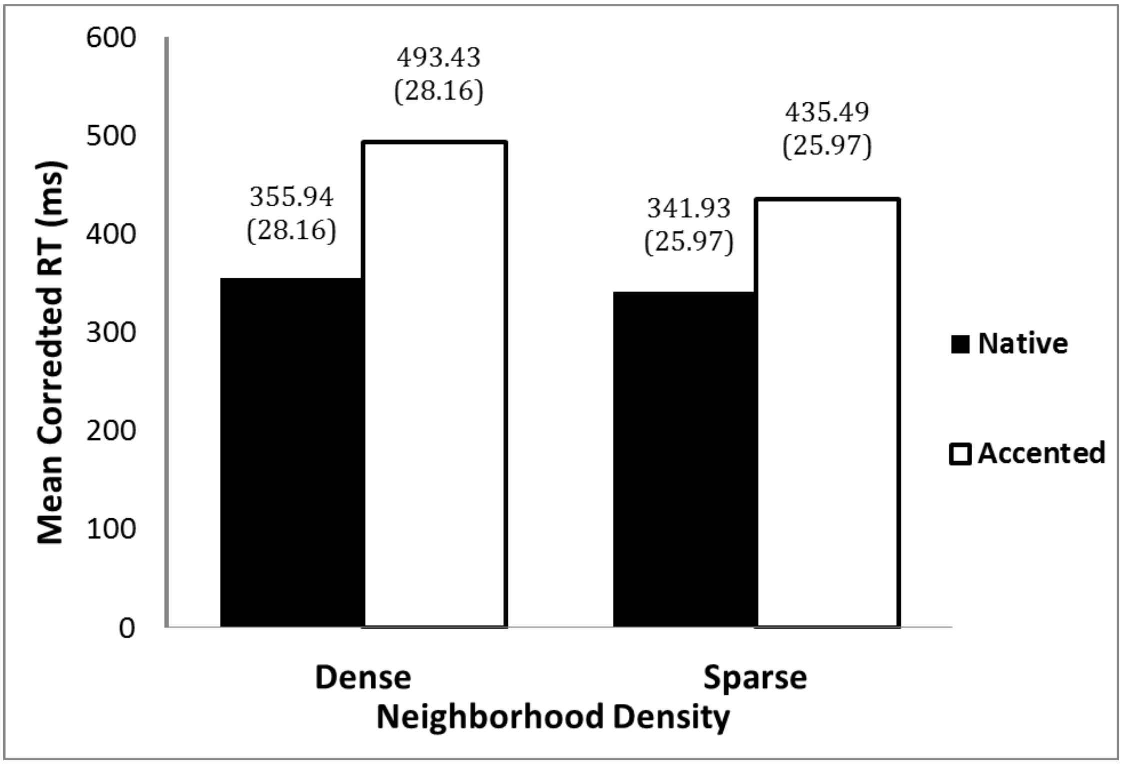

Figure 1. The graph presents the means corrected reaction times (S.E. in parentheses) for the lexical decision task in Experiment 1 as a function of accent type (native-produced vs. foreign-accented) and neighborhood density (dense vs, sparse; subject-analysis).

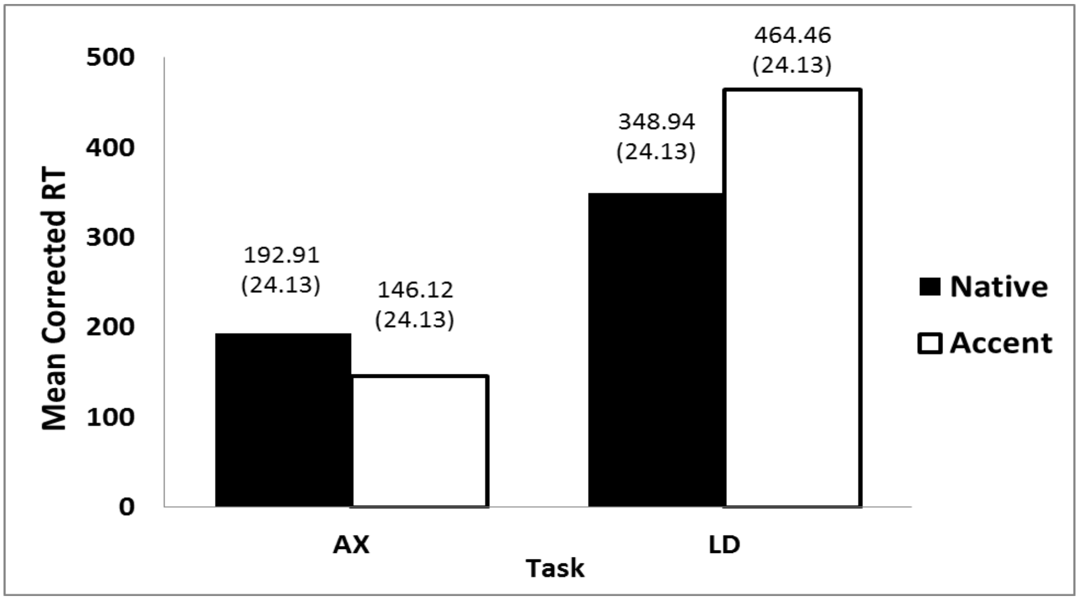

Figure 2a. The graph presents the means corrected reaction times (S.E. in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks in Experiments 1 and 2 respectively (subject-analysis).

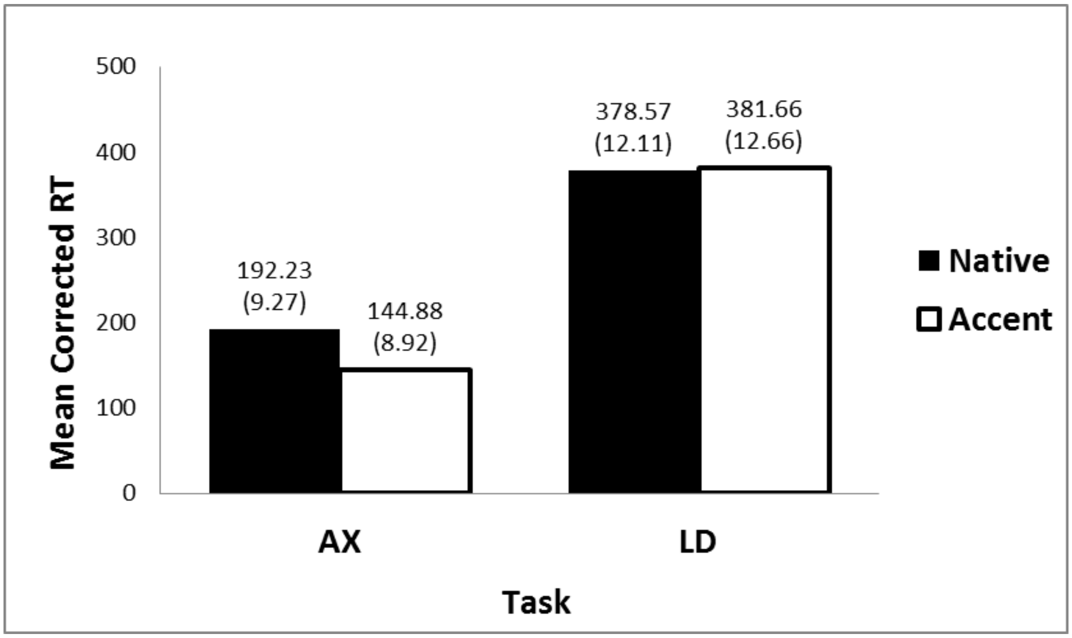

Figure 2b. The graph presents the means corrected reaction times (S.E. in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks in Experiments 1 and 2 respectively (item-analysis).

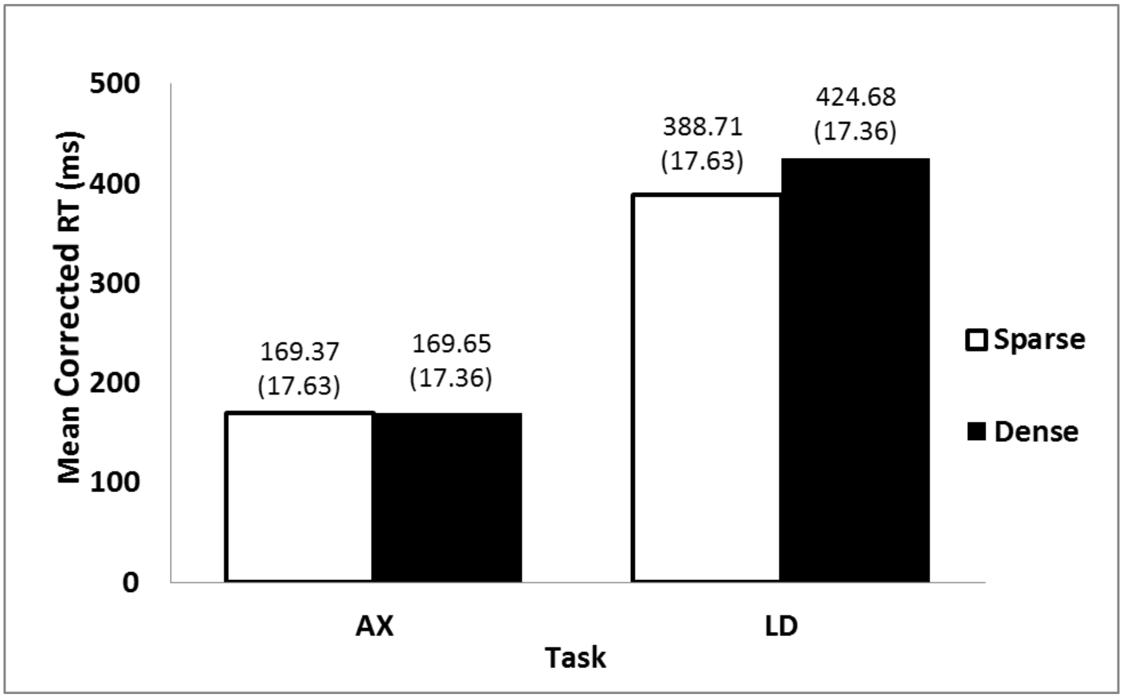

Figure 3. The graph presents the means corrected reaction times (S.E. in parentheses) for the dense and sparse words in the AX and LD tasks in Experiments 1 and 2 respectively (subject-analysis).

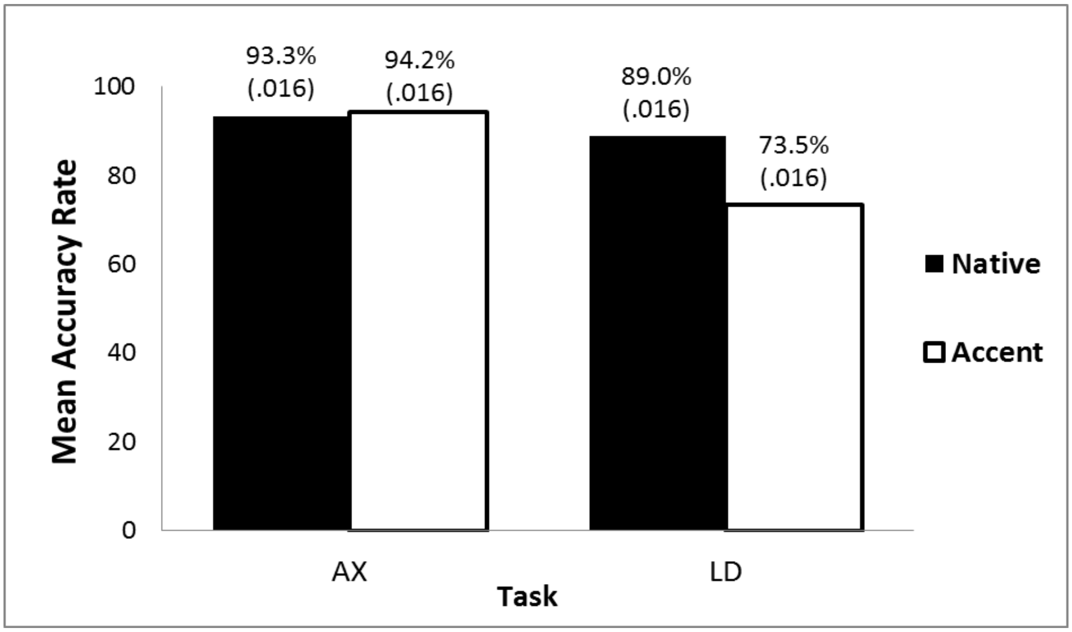

Figure 4a. The graph presents the means accuracy rate (S.E. in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks in Experiments 1 and 2 respectively (subject-analysis).

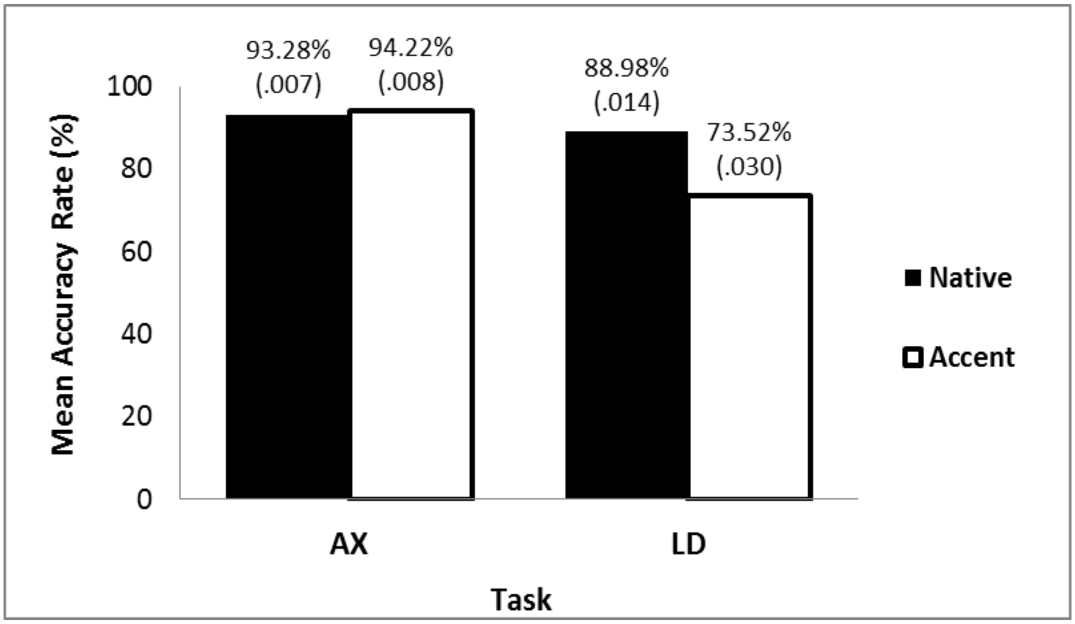

Figure 4b. The graph presents the means accuracy rate (S.E. in parentheses) for the foreign-accented and native-produced words in the AX and LD tasks in Experiments 1 and 2 respectively (item-analysis).

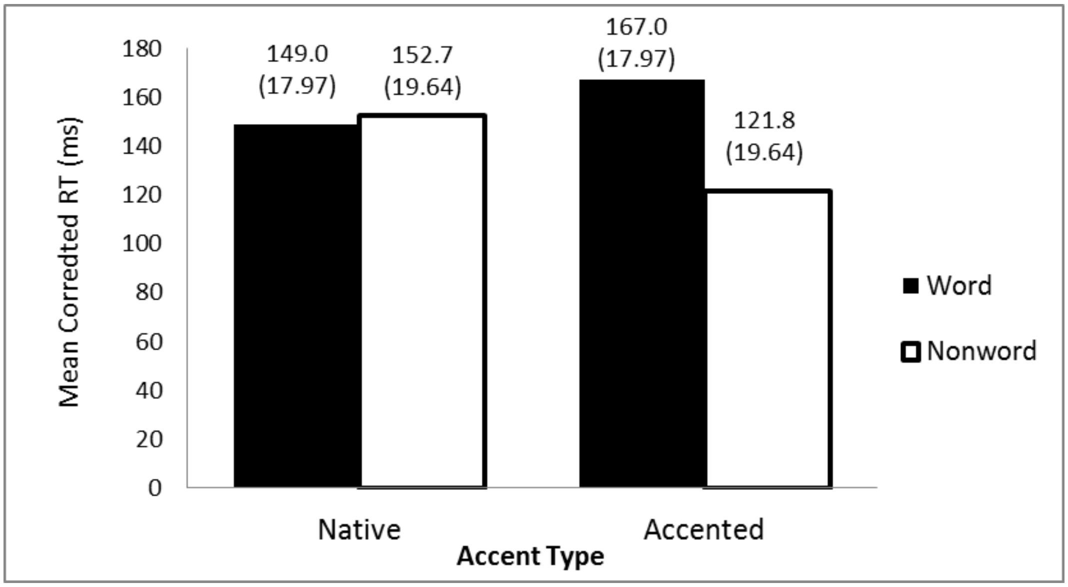

Figure 5a. The graph presents the means corrected reaction times (S.E. in parentheses) for words and nonwords in the foreign-accented and native-produced conditions in Experiment 3 (subject-analysis).

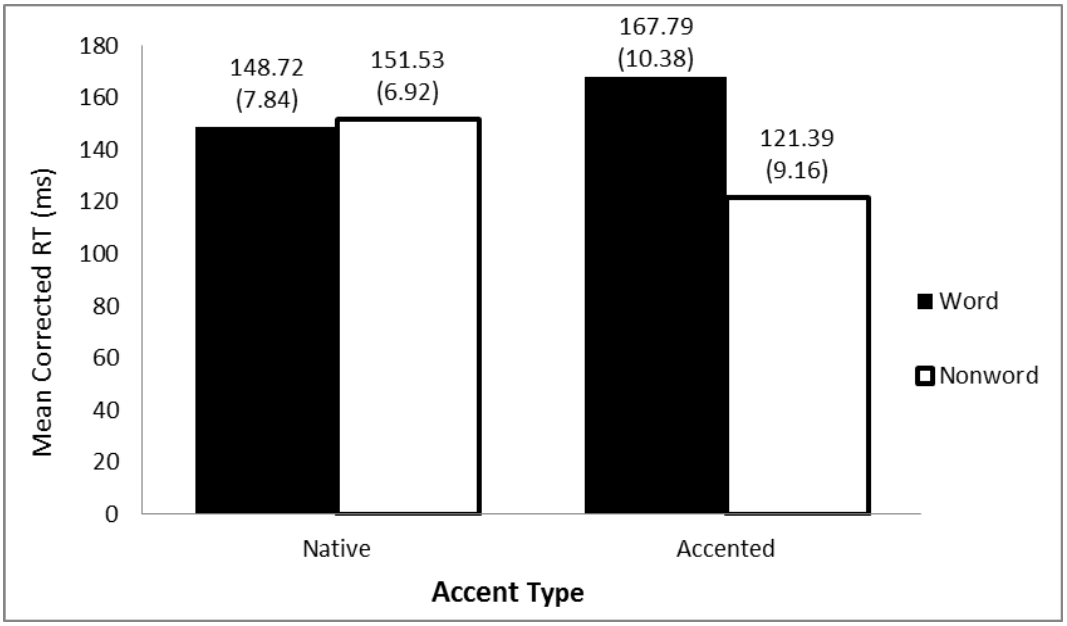

Figure 5b. The graph presents the means corrected reaction times (S.E. in parentheses) for words and nonwords in the foreign-accented and native-produced conditions in Experiment 3 (item-analysis).

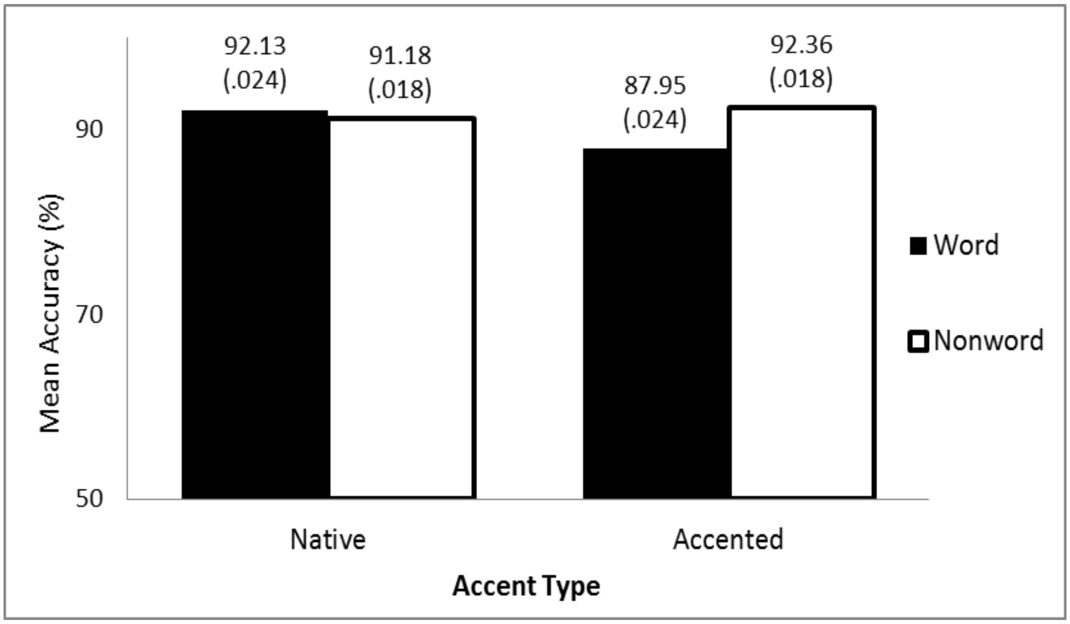

Figure 6a. The graph presents the mean accuracy rates (S.E. in parentheses) for the words and nonwords in the foreign-accented and native-produced conditions in Experiment 3 (subject-analysis).

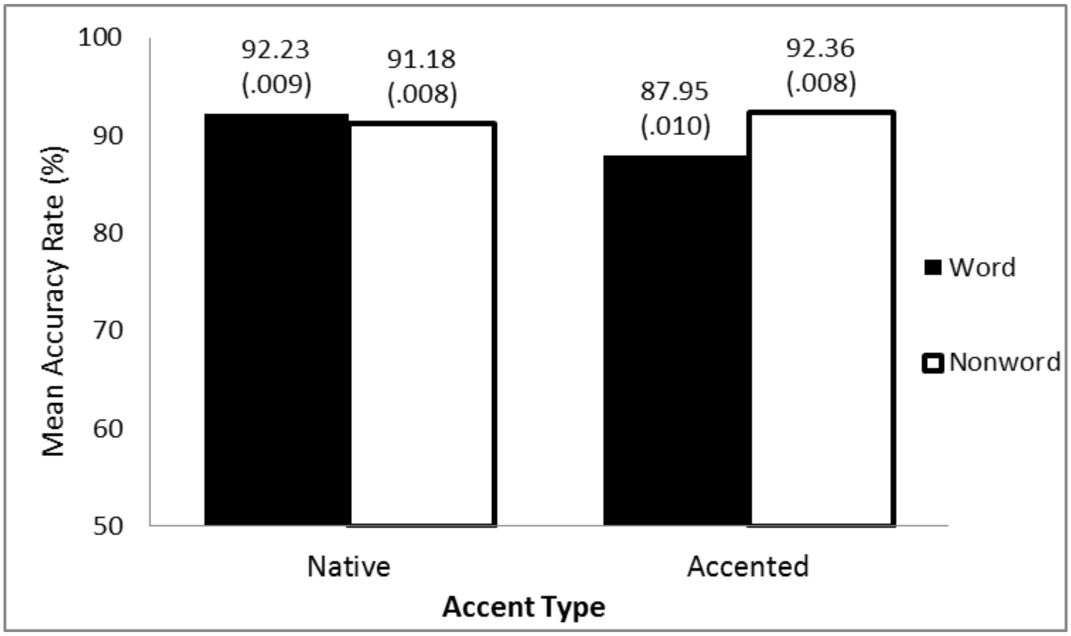

Figure 6b. The graph presents the mean accuracy rates (S.E. in parentheses) for the words and nonwords in the foreign-accented and native-produced conditions in Experiment 3 (item-analysis).

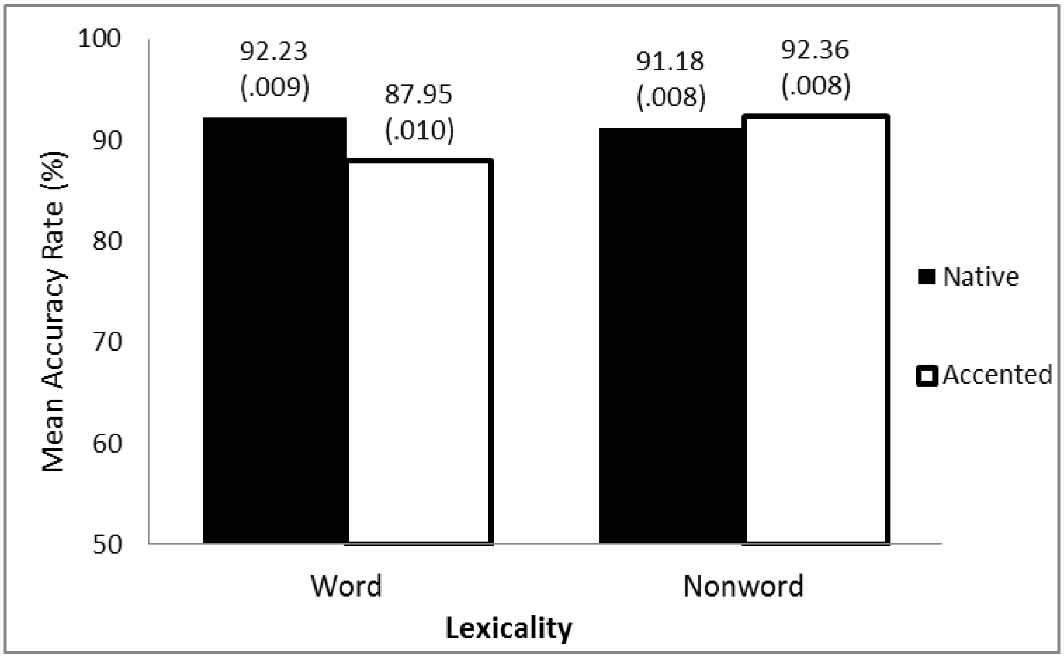

Figure 7. The graph presents the mean accuracy rates (S.E. in parentheses) for the native-produced and foreign-accented items in the two lexicality conditions in Experiment 3 (item-analysis).

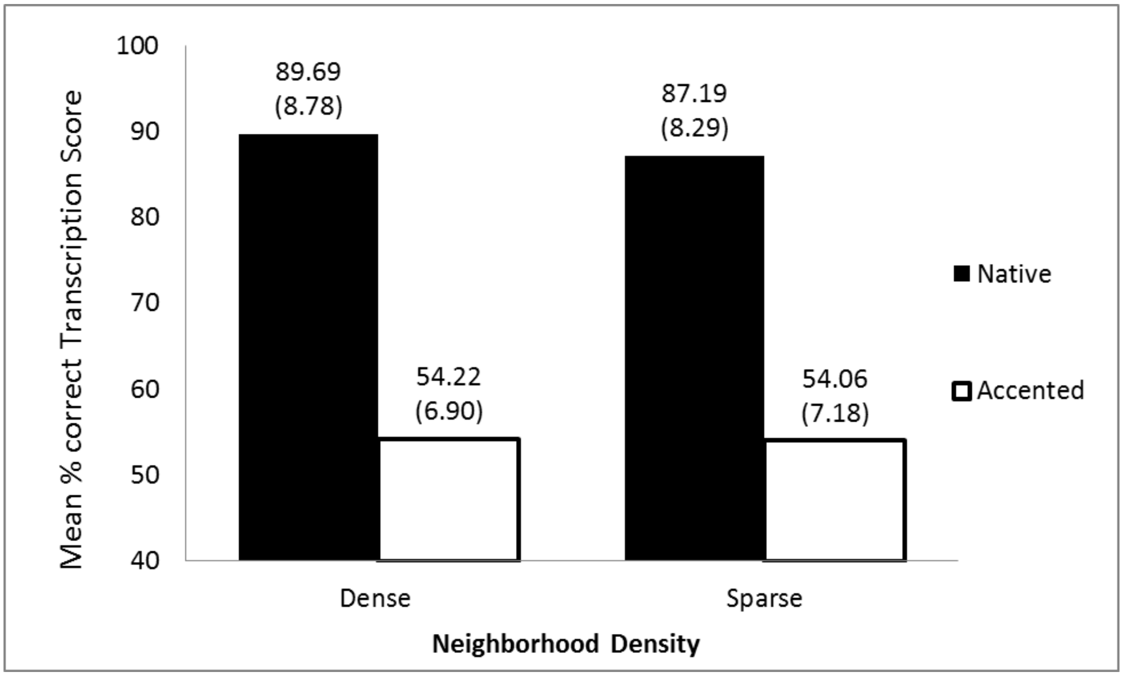

Figure 8. The graph presents the result of the perceptual identification task in Experiment 4— mean percent-correct transcription scores as a function of accent type (native-produced vs. foreign-accented) and neighborhood density (dense vs. sparse).

Table 1. Means accentedness ratings (standard deviations are in parentheses) for the sparse items, dense items and all items (overall rating) for each speaker.

	Accentedness Rating (1-7)		
Speaker	Sparse	Dense	Overall Rating
Native Male	1.3 (.7)	1.3 (.4)	1.3 (.6)
Native Female	1.8 (.9)	1.6 (.8)	1.7 (.9)
Accented Male	5.7 (1.0)	5.4 (1.1)	5.5 (1.1)
Accented Female	5.7 (.9)	5.7 (1.1)	5.7 (1.0)

Note. 1 = ”native-like”; 7 = ”strong foreign accent”

Table 2. Mean word durations in ms (standard deviations are in parentheses) for the word and nonword stimuli for each speaker and neighborhood density condition

	Mean Word Duration /ms (s.d.)					
	Words			Nonwords		
Speaker	Dense	Sparse	All words	Dense	Sparse	All Nonwords
Native Male	547.36 (84.27)	530.10 (73.44)	538.73 (78.89)	543.25 (91.51)	522.21 (78.39)	532.73 (85.19)
Native Female	644.05 (109.23)	645.44 (102.77)	644.74 (105.21)	677.59 (73.12)	682.82 (112.72)	680.21 (94.29)
Accented Male	601.50 (143.67)	603.67 (135.19)	602.58 (138.39)	567.54 (97.29)	557.83 (99.48)	562.69 (97.73)
Accented Female	680.73 (97.24)	716.82 (98.13)	698.77 (98.60)	669.49 (113.23)	661.36 (92.21)	665.43 (102.51)

Table 3. Mean corrected reaction time in *ms* and mean accuracy rate in *percentage* (standard deviations are in parentheses) as a function of neighborhood density (dense vs. sparse) and block of presentation (first vs. second) in Experiment 1.

	Neighborhood Density			
	Dense		Sparse	
Block	Corrected RTs	Accuracy Rates	Corrected RTs	Accuracy Rates
1st	528.53 (161.85)	75.31% (13.82%)	472.28 (136.10)	67.50% (14.28%)
2nd	456.26 (154.51)	75.00% (14.34%)	402.72 (155.12)	76.25% (10.84%)

Table 4. Descriptive Statistics for the Sub-lexical and Lexical Characteristics of the High and Low Phonotactic Probability Word Lists used in Experiment 3 and ANOVAs of the Mean Values.

High Phonotactic Probability Words	Positional Segment Frequency	Biphone Frequency	Neighborhood Density	Familiarity	log Word Freq.	Raw Freq.	Mean log Neighborhood Frequency
Mean	0.161	0.009	18.25	6.92	1.21	37.21	1.95
S.D.	0.014	0.003	4.024	0.213	0.594	50.38	0.222
Minimum	0.14	0.01	12.00	6.17	0.30	2.00	1.43
Maximum	0.19	0.02	25.00	7.00	2.25	177	2.25
Low Phonotactic Probability Words							
Mean	0.131	0.004	17.32	6.90	1.08	38.54	1.86
S.D.	0.016	0.001	5.313	0.157	0.656	80.89	0.287
Minimum	0.08	0.00	6.00	6.42	0.00	1.00	1.20
Maximum	0.16	0.01	25.00	7.00	2.59	391	2.38
ANOVAs							
F(1, 54)	56.18	59.76	0.544	0.089	0.562	0.005	1.728
p-value	< .0001	< .0001	0.464	0.767	0.457	0.942	0.194

Table 5. Descriptive Statistics for the Sub-lexical and Lexical Characteristics of the High and Low Phonotactic Probability Nonword Lists used in Experiment 3 and ANOVAs of the Mean Values.

High Phonotactic Probability Nonwords	Positional Segment Frequency	Biphone Frequency	Neighborhood Density	Mean log Neighborhood Frequency
Mean	0.136	0.0054	14.22	124.17
Standard deviation	0.013	0.0015	5.06	251.44
Minimum	0.117	0.0032	5	6.36
Maximum	0.167	0.0093	29	1221.17
Low Phonotactic Probability Nonwords				
Mean	0.111	0.0026	13.92	102.59
Standard deviation	0.011	0.0009	3.74	236.34
Minimum	0.093	0.0007	5	5.62
Maximum	0.129	0.0045	22	1222.33
ANOVAs				
F(1, 70)	80.692	97.039	0.085	0.141
p-value	< .0001	< .0001	0.772	0.709

Table 6. Means word duration in *ms* (standard deviations are in parentheses) for the word and nonword stimuli for each speaker and phonotactic probability (PP) condition in Experiment 3.

	Mean Word Duration /ms (s.d.)					
	Words			Nonwords		
Speaker	High PP	Low PP	All words	High PP	Low PP	All Nonwords
Native Male	537.16 (83.53)	518.34 (81.00)	527.75 (82.08)	574.74 (88.16)	561.73 (95.94)	568.24 (91.71)
Native Female	704.57 (67.59)	696.64 (78.64)	700.60 (72.77)	719.96 (81.25)	700.62 (75.37)	710.29 (78.42)
Accented Male	585.18 (104.47)	582.73 (97.69)	583.95 (100.22)	671.53 (113.29)	657.87 (122.97)	664.70 (117.59)
Accented Female	661.04 (93.11)	644.26 (98.17)	652.65 (95.18)	786.04 (119.88)	764.29 (98.17)	775.17 (109.34)

Table 7. The average segment and biphone probability for the word and nonword stimuli in the high and low phonotactic probability conditions and the corresponding magnitude differences between the two conditions in Vitevitch and Luce (1999), Vitevitch (2003) and Experiment 3 in the current study.

	High Probability Condition		**Low Probability Condition**		**Magnitude difference between the two conditions**	
	Avg. Segment Prob.	Avg. Biphone Prob.	Avg. Segment Prob.	Avg. Biphone Prob.	Avg. Segment Prob.	Avg. Biphone Prob.
Vitevitch & Luce (1999; Expt 1)						
Nonwords	.1926	0.0143	0.543	0.0006	.3504	0.0137
Words	0.201	0.0123	0.126	0.0048	0.075	0.0075
Vitevitch (2003)						
Words	0.203	0.0140	0.135	0.006	0.068	0.0080
Current Experiment						
Nonwords	0.136	0.0054	0.111	0.0026	0.025	0.0029
Words	0.161	0.0090	0.131	0.0040	0.030	0.0050

Table 8. Means Percent Correct Transcription Scores (standard deviations are in parentheses) for each of the two levels of neighborhood density (dense, sparse) for each of the four speakers in Experiment 4.

	Mean Percent Correct Transcription Scores			
ND Condition	NMS	NFS	AMS	AFS
Dense	86.56 (11.32)	92.81 (3.62)	55.00 (8.23)	53.43 (5.60)
Sparse	84.06 (9.25)	90.31 (6.15)	51.56 (6.79)	56.56 (6.98)

Note. NMS (native male speaker), NFS (native female speaker), AMS (accented male speaker), and AFS (accented female speaker).

Table 9. The observed frequency counts of misperceptions from Experiment 4 in a 2 (accent type: native vs. foreign-accented) x 2 (neighborhood density: dense vs. spare) x 2 (phonological similarity with target word: low vs. similar) contingency table.

Accented	**High Similarity**	**Low Similarity**
Dense	242	51
Sparse	175	119
Native		
Dense	51	15
Sparse	36	46

Table 10. The observed frequency counts (conditional proportions in parentheses) of the misperceptions from Experiment 4 in a 2 (accent type: native vs. foreign-accented) x 2 (consistency: consistent vs. not consistent) contingency table.

	Consistent	**Not Consistent**
Native	81 (55.9%)	64 (44.1%)
Accent	476 (82.8%)	99 (17.2%)

Table 11. Results of fitting several log-linear models to the misperception data from Experiment 4 using a backward elimination procedure with the likelihood-ratio model comparison test.

Model	**Predictors**	**Chi-Square (G^2)**	***DF***	**Models Compared**	**Difference in G^2**	**Difference in *DF***	***p*-value**
M1	AT x ND x PS	0	0	---	---	---	---
M2	AT x ND, AT x PS, ND x PS	0.52	1	M2 - M1	0.52	1	0.47
M3a*	AT x PS, ND x PS	0.67	2	M3a - M2	0.15	1	0.70
M3b	AT x ND, ND x PS	7.29	2	M3b - M2	6.77	1	0.01
M3c	AT x ND, AT x PS	56.17	2	M3c - M2	55.65	1	0.00
M4a	ND x PS	8.63	3	M4a - M3a	7.962	1	0.01
M4b	AT x ND	57.51	3	M4b - M3a	56.841	1	0.00

Note. *DF* = degree of freedom, AT = Accent Type, ND = Neighborhood Density, PS = Phonological Similarity

*Final model after backward elimination

Table 12. Summary of the Four Experimental Designs.

Expt.	Task	Processing Emphasized	Independent Variables	Dependent Variables	Stimuli
1	Lexical decision	Lexical	Neighborhood density, Accent type	Reaction time, Accuracy rate	Words
2	AX	Sub-lexical	Neighborhood density, Accent type, task	Reaction time, Accuracy rate	Words from Expt. 1
3	AX	Sub-lexical	Phonotactic probability, Accent type	Reaction time, Accuracy rate	Words and nonwords
4	Perceptual identification	Lexical	Neighborhood density, Accent type	Transcription accuracy, Misperceptions type	Words from Expt. 1

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The Relationship between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentals, Prosody, and Syllable Structure. *Language Learning, 42*(4), 529-555.

Bartkova, K., & Jouvet, D. (2007). On using units trained on foreign data for improved multiple accent speech recognition. *Speech Communication, 49*(10-11), 836-846.

Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvet, D., . . . Wellekens, C. (2007). Automatic speech recognition and speech variability: A review. *Speech Communication, 49*(10-11), 763-786. doi: 10.1016/j.specom.2007.02.006

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05) [Computer program]. (Version Retrieved May 1, 2009, from http://www.praat.org/).

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729.

Brennan, E. M., & Brennan, J. S. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech, 24*(3), 207-221.

Burki-Cohen, J., Miller, J. L., & Eimas, P. D. (2001). Perceiving Non-Native Speech. *Language and Speech, 44*(2), 149-169.

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

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning & Verbal Behavior. Vol., 12*(4), 335-359.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America, 116*(6), 3647-3658.

Delamere, T. (1996). The Importance of Interlanguage Errors with Respect to Stereotyping by Native Speakers in Their Judgements of Second Language Learners' Performance. *System, 24*(3), 279-297.

Derwing, T. M., & Munro, M. J. (1997). Accent, Intelligibility, and Comprehensibility: Evidence from Four L1s. *Studies in Second Language Acquisition, 19*(1), 1-16.

Eimas, P. D., Hornstein, S. M., & Payton, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory and Language, 29*(2), 160-180.

Flege, J. E. (1984). The detection of French accent by American listeners. *The Journal of the Acoustical Society of America, 76*(3), 692-707.

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

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

Gill, M. M. (1994). Accent and stereotypes: Their effect on perceptions of teachers and lecture comprehension. *Journal of Applied Communication Research, 22*(4), 348-361.

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

Graddol, D. (Ed.). (1997). *The future of English? A guide to forecasting the popularity of the English language in the 21st century.* London (United Kingdom): British Council.

Greenberg, J. H., & Jenkins, J. J. (1967). Studies in the psychological correlates of the sound system of American English. In L. A. Jakobovits & M. S. Miron (Eds.), *Readings in the psychology of language* (pp. xi, 636 p.186-200). Englewood Cliffs, N.J.,: Prentice-Hall.

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical Frequency and Neighborhood Density Effects on the Recognition of Native and Spanish-Accented Words by Native English and Spanish Listeners. *The Journal of the Acoustical Society of America, 117*(2), 896-907.

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, Comprehensibility, and Accentedness of L2 Speech: The Role of Listener Experience and Semantic

Context. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 64*(3), 459-489.

Koster, C. J., & Koet, T. (1993). The Evaluation of Accent in the English of Dutchmen. *Language Learning, 43*(1), 69-92.

Kučera, H., & Francis, W. N. (Eds.). (1967). *Computational Analysis of Present Day American English.* Providence: Brown University Press.

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*(3), 245-294.

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning & Verbal Behavior. Vol., 12*(2), 119-131.

Lane, H. (1963). Foreign accent and speech distortion. *Journal of the Acoustical Society of America, 35*(4), 451-453.

Lawson, A. D., Harris, D. M., & Grieco, J. J. (2003). *Effect of foreign accent on speech recognition in the NATO n-4 corpus*. Paper presented at the EUROSPEECH-2003, Geneva, Switzerland.

Levi, S. V., Winters, S. J., & Pisoni, D. B. (2007). Speaker-Independent Factors Affecting the Perception of Foreign Accent in a Second Language. *The Journal of the Acoustical Society of America, 121*(4), 2327-2338.

Lewandowsky, S. (1993). The rewards and hazards of computer simulations. *Psychological Science, 4*(4), 236-243.

Lippmann, R. P. (1997). Speech recognition by machines and humans. *Speech Communication, 22*(1), 1-15. doi: 10.1016/s0167-6393(97)00021-6

Luce, P. A., Goldinger, S. D., Auer, E. T., Jr., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics, 62*(3), 615-625.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing, 19*(1), 1-36.

Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics, 26*(4), 381-400.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1-86.

Moore, R. K. (2007). Spoken language processing: Piecing together the puzzle. *Speech Communication. Special Issue: Bridging the gap between human and automatic speech recognition, 49*(5), 418-435. doi: 10.1016/j.specom.2007.01.011

Munro, M. J. (1993). Productions of English Vowels by Native Speakers of Arabic: Acoustic Measurements and Accentedness Ratings. *Language and Speech, 36*(1), 39-66.

Munro, M. J. (1998). The Effects of Noise on the Intelligibility of Foreign-Accented Speech. *Studies in Second Language Acquisition, 20*(2), 139-154.

Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306.

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48*(2), 159-182.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*(1), 111-131.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*(3), 299-370.

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, 10*, 357-376.

Owren, M. J. (2008). GSU Praat tools: Scripts for modifying and analyzing sounds using Praat acoustics software. *Behavior Research Methods, 40*(3), 822-829. doi: 10.3758/brm.40.3.822

Raaijmakers, J. G. W. (2003). A further look at the "language-as-fixed-effect fallacy". *Canadian Journal of Experimental Psychology, 57*(3), 141-151.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*(3), 416-426.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 55A*(4), 1339-1362.

Reed, M. (2000). He Who Hesitates: Hesitation Phenomena as Quality Control in Speech Production, Obstacles in Non-Native Speech Perception. *Journal of Education, 182*(3), 67-91.

Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2005). Release from Information Making by Time Reversal of Native and Non-Native Interfering Speech. *The Journal of the Acoustical Society of America, 118*(3), 1274-1277.

Riazantseva, A. (2001). Second Language Proficiency and Pausing: A Study of Russian Speakers of English. *Studies in Second Language Acquisition, 23*(4), 497-526.

Ryan, E. B., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and Speech, 20*(3), 267-273.

Schairer, K. E. (1992). Native Speaker Reaction to Non-Native Speech. *The Modern Language Journal, 76*(3), 309-319.

Scharenborg, O. (2007). Reaching over the gap: A review of efforts to link human and automatic speech recognition research. *Speech Communication. Special Issue: Bridging the gap between human and automatic speech recognition, 49*(5), 336-347. doi: 10.1016/j.specom.2007.01.009

Scharenborg, O., Norris, D., ten Bosch, L., & McQueen, J. M. (2005). How Should a Speech Recognizer Work? *Cognitive Science: A Multidisciplinary Journal, 29*(6), 867-918. doi: 10.1207/s15516709cog0000_37

Schmid, P. M., & Yeni-Komshian, G. H. (1999). The effects of speaker accent and target predictability on perception of mispronunciations. *Journal of Speech, Language, and Hearing Research, 42*(1), 56-64.

Sidaras, S. K., Alexander, J. E. D., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America, 125*(5), 3306-3316.

Stafford, D. (2009, Aug. 10, 2009). Outsourcing agreements continue decade-long acceleration, *The Kansas City Star*. Retrieved from http://www.kansascity.com/2009/08/10/1374748/outsourcing-agreements-continue.html

Storkel, H. L., & Hoover, J. R. (2010). An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. *Behavior Research Methods, 42*, 497-506.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662. doi: 10.1037/h0054651

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25*(1), 1-24.

Tarone, E. E. (1980). Some Influences on the Syllable Structure of Interlanguage Phonology. *IRAL, International Review of Applied Linguistics in Language Teaching, 18*(2), 139-152.

Temple, L. (2000). Second Language Learner Speech Production. *Studia Linguistica, 54*(2), 288-297.

ten Bosch, L., & Kirchhoff, K. (2007). Bridging the gap between human and automatic speech recognition. *Speech Communication. Special Issue: Bridging the gap between human and automatic speech recognition, 49*(5), 331-335. doi: 10.1016/j.specom.2007.03.001

Van Compernolle, D. (2001). Recognizing speech of goats, wolves, sheep and â€¦ non-natives. *Speech Communication, 35*(1-2), 71-79.

Van Engen, K. J. (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. *Speech Communication. Special Issue: Non-native speech perception in adverse conditions, 52*(11-12), 943-953. doi: 10.1016/j.specom.2010.05.002

Van Engen, K. J., & Bradlow, A. R. (2007). Sentence Recognition in Native- and Foreign-Language Multi-Talker Background Noise. *The Journal of the Acoustical Society of America, 121*(1), 519-526.

Van Wijngaarden, S. J. (2001). Intelligibility of Native and Non-Native Dutch Speech. *Speech Communication, 35*, 1-2.

Vitevitch, M. S. (2002). Naturalistic and Experimental Analyses of Word Frequency and Neighborhood Density Effects in Slips of the Ear. *Language and Speech, 45*(4), 407-434.

Vitevitch, M. S. (2003). The influence of sublexical and lexical representations on the processing of spoken words in English. *Clinical Linguistics & Phonetics, 17*(6), 487-499.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9*(4), 325-329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*(3), 374-408.

Vitevitch, M. S., & Rodriguez, E. (2004). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders, 3*, 64-73.

Vitevitch, M. S., Stamer, M. K., & Sereno, J. A. (2008). Word length and lexical competition: Longer is the same as shorter. *Language and Speech, 51*(4), 361-383.

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of Acoustic Variability in the Perceptual Learning of Non-Native-Accented Speech Sounds. *Phonetica, 64*, 2-3.

You, H., Alwan, A., Kazemzadeh, A., & Narayanan, S. (2005). *Pronunciation variations of Spanish-accented English spoken by young children.* Paper presented at the Interspeech 2005- Eurospeech, Lisboa, Portugal.

Zechner, K., Higgins, D., Xi, X., & Williamson, D. M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication. Special Issue: Spoken language technology for education, 51*(10), 883-895. doi: 10.1016/j.specom.2009.04.009

APPENDIX A.1- Dense Neighborhood Density Words used in Experiment 1

Stimulus Word	**Neighborhood Density**	**Familiarity**	**log Word Freq.**	**Pos. Seg. Freq.**	**Biphone Freq.**	**log NF**
bug	26	7.00	0.60	0.1083	0.0047	1.80
buck	29	7.00	1.30	0.1439	0.0053	1.83
bought	25	7.00	1.75	0.1337	0.0025	2.34
duck	25	6.75	0.95	0.1445	0.0043	1.81
dumb	29	7.00	1.11	0.1404	0.0075	2.01
dune	27	7.00	0.00	0.17	0.0043	1.96
dine	30	7.00	0.30	0.1822	0.0082	2.04
fun	25	7.00	1.64	0.1819	0.0067	2.06
fall	26	7.00	2.17	0.1368	0.005	2.45
fine	28	6.92	2.21	0.177	0.0065	2.12
cop	30	7.00	1.18	0.1903	0.0191	1.84
call	26	7.00	2.27	0.1829	0.006	2.16
lash	26	6.17	0.78	0.1212	0.0065	1.63
lock	31	7.00	1.36	0.1481	0.0052	1.93
lead	31	7.00	2.42	0.145	0.0077	2.10
lease	27	6.92	1.00	0.1447	0.0042	2.02
leave	26	7.00	2.31	0.0895	0.0038	1.76
kneel	27	7.00	0.70	0.1293	0.0044	1.74
nip	25	7.00	0.48	0.1571	0.0068	1.66
pop	29	7.00	0.90	0.182	0.0103	1.69
rash	26	6.58	0.00	0.1372	0.007	1.63
raise	30	7.00	1.72	0.0994	0.0042	1.86
son	26	7.00	2.44	0.2377	0.0116	2.43
seek	31	6.92	1.84	0.1877	0.005	2.07
shear	26	7.00	1.74	0.1843	0.0062	2.19
shore	28	7.00	1.79	0.1374	0.0188	2.60
tuck	28	6.83	0.30	0.1372	0.0033	1.95
tall	27	7.00	1.74	0.1347	0.0044	2.25
tune	27	7.00	1.00	0.1627	0.0047	52.25
wed	25	7.00	0.30	0.1312	0.0061	2.37
wick	26	6.67	0.60	0.17	0.0132	2.41
wine	30	7.00	1.86	0.1507	0.0064	1.98

Note: **Pos. Seg. Freq.** is position segment frequency (a measure of phonotactic probability); **Biphone Freq.** is biphone frequency (a measure of phonotactic probability). **NF** is neighborhood frequency;

APPENDIX A.2- Sparse Neighborhood Density Words used in Experiment 1

Stimulus Word	Neighborhood Density	Familiarity	log Word Freq.	Pos. Seg. Freq.	Biphone Freq.	log NF
buzz	15	7.00	1.11	0.1105	0.0044	1.92
bib	13	6.83	0.30	0.1734	0.0064	2.25
beam	16	6.92	1.32	0.1324	0.0034	2.19
dash	15	6.92	1.04	0.1389	0.0039	1.52
dawn	19	7.00	1.45	0.1644	0.0022	2.10
deed	18	7.00	0.90	0.1216	0.0052	2.33
deep	18	7.00	2.04	0.1207	0.0045	1.97
fad	19	6.33	0.30	0.164	0.0058	2.21
far	18	6.58	2.63	0.1855	0.018	2.43
fish	13	7.00	1.54	0.1505	0.0059	1.87
gone	17	7.00	2.29	0.1386	0.002	1.81
cup	18	7.00	1.65	0.169	0.0055	2.07
calm	17	7.00	1.54	0.2026	0.0224	1.95
kiss	13	7.00	1.23	0.2677	0.0188	2.34
lull	15	6.25	0.30	0.147	0.0064	1.66
love	11	6.67	2.37	0.0969	0.003	1.91
lawn	19	7.00	1.18	0.1467	0.003	2.41
league	19	7.00	1.84	0.0838	0.003	1.86
null	17	6.17	1.11	0.1367	0.006	1.51
neck	13	7.00	1.91	0.1502	0.0094	1.74
pool	18	7.00	2.05	0.1802	0.0018	1.99
wreck	18	7.00	0.90	0.1765	0.0156	1.91
robe	18	7.00	0.78	0.1254	0.0039	1.94
shun	19	6.33	0.00	0.145	0.0062	2.24
psalm	11	6.92	0.60	0.2123	0.008	2.27
chute	17	7.00	1.46	0.0978	0.0029	1.97
sour	10	6.92	0.48	0.1905	0.0009	1.76
tug	18	7.00	0.48	0.1016	0.0027	1.71
tape	16	7.00	1.54	0.1108	0.0029	2.08
town	14	7.00	2.33	0.1503	0.0045	2.19
walk	15	7.00	2.00	0.0903	0.003	2.20
wipe	14	7.00	1.00	0.0917	0.003	2.19

Note: **Pos. Seg. Freq.** is position segment frequency (a measure of phonotactic probability); **Biphone Freq.** is biphone frequency (a measure of phonotactic probability). **NF** is neighborhood frequency;

APPENDIX B- International Phonetic Alphabet (IPA) Transcriptions of the Nonword Foils used in Experiment 1

bʌtʃ
bʌp
bɒf
dʌʃ
dʌt
duθ
daɪp
fʌm
foθ
faɪb
kɒz
kub
læt
lɒd
lɛm
lið
lim
niv
nɪθ
pɒg
ræd
rem
sʌv
sig
ʃɪk
ʃof
tʌdʒ
tɒb
tuv
wɛg
wɪd
waɪm
bʌb
bɪdʒ
bil
dæz
duf
dik
dit
fæf
fɒn
fɪd
guð
kʌk
kɒŋ
kɪg
lʌt
lʌθ
lok
lib
nʌs
nɛd
puk
rɛl
rof
ʃʌŋ
sɒg
ʃul
saʊt
tʌl
ten
taʊs
wuk
waɪb

APPENDIX C- International Phonetic Alphabet (IPA) Transcriptions of the nonword foils used for SAME pairs in Experiment 2

bɛtʃ
tʃæk
tʃæs
tʃʌl
tʃɒm
dek
dɛʃ
dɝs
dul
dut
fætʃ
fob
dʒɪv
kʌʃ
kʌz
koθ
lɛz
mʌv
mʌz
mɒtʃ
mɒdʒ
nos
rædʒ
rɛʃ
sʌʃ
ʃɒl
ʃɛm
sɝg
sɝp
sɝθ
tæθ
θʌl

vɪʃ
vɪθ
zin
zaɪn
tʃæl
dis
fɝs
kʌv
kɒtʃ
kɒθ
kof
mɛtʃ
mɛz
mɝn
maʊn
pɒf
pɒdʒ
pɒθ
pim
pob
pof
pog
ʃæs
sɛʃ
θæl
θæs
θɒn
vɛd
zæl
zæs
zɪt
biv

tʃɛf
tʃɛg
tʃop
tʃut
dʌdʒ
dɒʃ
dog
doθ
dɝm
fʌf
fɛʃ
fɝv
gek
dʒaɪl
kɝtʃ
kɝθ
lʌz
mʌdʒ
mʌθ
mum
næz
nʌp
rʌz
rɒg
ʃɒb
ʃɛb
saʊʃ
θæʃ
taʊl
vætʃ
væʃ
jɝl

zɪtʃ
zɪθ
zun
dotʃ
fɝg
gɒf
gɛb
gɛg
giz
gop
gaʊl
lɒʃ
nædʒ
næθ
nɝt
ʃiv
θɒb
θɒg
θod
vaʊs
wætʃ
wæθ
wʌk
wʌp
wɒb
wɒf
zæf
zɛb
zɛg
bæb
bɔɪn
baʊn

tʃæz
tʃɛd
tʃɪθ
dɝd
daʊd
fʌtʃ
fɒŋ
fuf
gɒg
gin
gaɪn
dʒʌp
dʒæŋ
dʒuf
kɛp
kaɪs
kim
lɝg
leb
ludʒ
mɪg
mɔɪd
mɝd
næs
nɒf
nik
nɝm
pæʃ
pʌdʒ
pɪb
pɝp
rʌdʒ

riʃ
rɝd
ʃæb
ʃʌd
sæf
sɒʃ
sɛs
sɪdʒ
θæg
θʌp
θɒf
tip
tɪθ
tɔɪd
tup
væp
vɒm
vep
við
vɪdʒ
wʌg
wɛm
wig
wof
jæs
jɒŋ
jib
jum
zæŋ
zɒk
zɛs
zuv

APPENDIX D- International Phonetic Alphabet (IPA) Transcriptions of the nonword foils used for forming the DIFFERENT pairs in Experiment 2.

bʌtʃ
bʌp
bɒf
dʌʃ
dʌt
duθ
daɪp
fʌm
foθ
faɪb
kɒz
kub
læt
lɒd
lɛm
lið
lim
niv
nɪθ
pɒg
ræd
rem
sʌv
sig
ʃɪk
ʃof
tʌdʒ
tɒb
tuv
wɛg
wɪd
waɪm

bʌb
bɪdʒ
bil
dæz
duf
dik
dit
fæf
fɒn
fɪd
guð
kʌk
kɒŋ
kɪg
lʌt
lʌθ
lok
lib
nʌs
nɛd
puk
rɛl
rof
ʃʌŋ
sɒg
ʃul
saʊt
tʌl
ten
taʊs
wuk
waɪb

bɛm
tʃæm
tʃæg
tʃʌv
tʃɒʃ
dev
dɛp
dɝg
duð
dub
fæʃ
fof
dʒɪtʃ
kʌg
kʌθ
kob
lɛb
mʌb
mʌŋ
mɒn
mɒʃ
nog
ræf
rɛm
sʌg
ʃɒf
ʃɛtʃ
sɝm
sɝb
sɝz
tæl
θʌʃ

vɪf
vɪk
zif
zaɪt
tʃæv
dib
fɝb
kʌn
kɒŋ
kɒdʒ
koz
mɛk
mɛl
mɝs
maʊt
pɒb
pɒtʃ
pɒz
pidʒ
pom
poθ
pov
ʃæv
sɛf
θæd
θæp
θɒm
vɛl
zæʃ
zæv
zɪm
bið

tʃɛp
tʃɛʃ
tʃof
tʃus
dʌp
dɒŋ
dof
dod
dɝp
fʌv
fɛt
fɝp
ges
dʒaɪk
kɝf
kɝg
lʌp
mʌn
mʌp
mudʒ
næd
nʌd
rʌtʃ
rɒp
ʃɒŋ
ʃɛn
saʊdʒ
θæv
taʊz
væθ
væg
jɝm

zɪs
zɪg
zuk
dok
fɝdʒ
gɒk
gɛd
gɛl
gif
goθ
gaʊθ
lɒm
næf
næl
nɝg
ʃid
θɒk
θɒp
θov
vaʊl
wæp
wæz
wʌm
wʌʃ
wɒŋ
wɒθ
zæb
zɛtʃ
zɛk
bæf
bɔɪs
baʊs

tʃæʃ
tʃɛn
tʃɪv
dɝn
daʊθ
fʌŋ
fɒk
fuk
gɒm
git
gaɪt
dʒʌs
dʒædʒ
dʒuv
kɛt
kaɪv
kiv
lɝs
lef
luθ
mɪp
mɔɪn
mɝm
næm
nɒθ
niʃ
nɝz
pæv
pʌθ
pɪf
pɝθ
rʌs

rit
rɝs
ʃæp
ʃʌz
sæθ
sɒθ
sɛg
sɪʃ
θæf
θʌn
θɒθ
tiv
tɪd
tɔɪs
tuð
væl
vɒt
vev
vim
vɪs
wʌz
wɛv
wim
wos
jæl
jɒp
jif
jut
zæn
zɒθ
zɛl
zup

APPENDIX E.1- High Phonotactic Probability Words used in Experiment 3

Stimulus Word	Neighborhood Density	Familiarity	log Word Freq.	Raw Freq.	log NF	Pos. Seg. Freq.	Biphone Freq.
badge	13	6.9	0.70	5	2.04	0.1414	0.0067
bus	20	7.0	1.54	35	2.06	0.1692	0.0073
balm	13	7.0	1.56	36	1.84	0.1611	0.0097
bed	25	7.0	2.10	127	2.23	0.1621	0.0069
dull	23	7.0	1.43	27	1.81	0.1647	0.0070
dock	22	7.0	1.45	28	1.83	0.1658	0.0057
doll	16	6.9	1.00	10	1.87	0.1860	0.0082
dead	24	7.0	2.24	174	2.25	0.1627	0.0108
deaf	13	7.0	1.08	12	1.97	0.1444	0.0086
dish	12	7.0	1.20	16	2.22	0.1557	0.0164
fad	19	6.3	0.30	2	2.21	0.1640	0.0058
foam	16	6.9	1.57	37	2.24	0.1453	0.0084
gun	20	7.0	2.07	118	2.16	0.1613	0.0073
cash	25	7.0	1.57	37	1.61	0.1798	0.0142
coach	14	7.0	1.38	24	2.00	0.1500	0.0066
cove	18	7.0	0.30	2	1.88	0.1656	0.0074
lid	23	7.0	1.28	19	2.06	0.1683	0.0094
live	15	7.0	2.25	177	1.94	0.1539	0.0093
match	14	7.0	1.61	41	2.23	0.1446	0.0116
math	15	7.0	0.60	4	2.23	0.1440	0.0111
mob	15	7.0	1.00	10	1.65	0.1437	0.0083
mop	16	7.0	0.48	3	1.80	0.1548	0.0089
pod	23	6.2	0.48	3	1.83	0.1829	0.0103
wreck	18	7.0	0.90	8	1.91	0.1765	0.0156
rich	21	7.0	1.87	74	1.82	0.1543	0.0184
sash	20	6.5	0.48	3	1.63	0.1895	0.0066
sub	17	7.0	0.70	5	1.83	0.1676	0.0095
tag	21	7.0	0.70	5	1.43	0.1418	0.0067

Note: **Pos. Seg. Freq.** is position segment frequency (a measure of phonotactic probability); **Biphone Freq.** is biphone frequency (a measure of phonotactic probability). **NF** is neighborhood frequency;

APPENDIX E.2- Low Phonotactic Probability Words used in Experiment 3

Stimulus Word	Neighborhood Density	Familiarity	log Word Freq.	Raw Freq.	log NF	Pos. Seg. Freq.	Biphone Freq.
bud	23	6.8	0.95	9	1.98	0.1284	0.0044
beg	14	7.0	1.04	11	2.03	0.1420	0.0048
berth	16	7.0	1.85	70	1.93	0.0833	0.0022
dash	15	6.9	1.04	11	1.52	0.1389	0.0039
dug	22	7.0	1.20	16	1.83	0.1089	0.0037
duck	25	6.8	0.95	9	1.81	0.1445	0.0043
dodge	8	6.8	1.04	11	1.81	0.1231	0.0032
dome	25	7.0	1.23	17	1.69	0.1505	0.0036
dove	16	7.0	0.60	4	1.63	0.1247	0.0024
fetch	9	7.0	0.78	6	1.65	0.1275	0.0031
fame	24	7.0	1.26	18	2.38	0.1252	0.0035
gab	17	6.7	0.00	1	1.47	0.1314	0.0058
give	7	7.0	2.59	391	1.92	0.1458	0.0043
cook	15	7.0	1.67	47	2.17	0.1564	0.0015
latch	18	7.0	0.70	5	1.83	0.1215	0.0060
laugh	19	7.0	1.45	28	1.84	0.1332	0.0058
mud	20	7.0	1.51	32	1.63	0.1344	0.0049
muff	18	6.4	0.00	1	1.52	0.1161	0.0055
mesh	6	6.6	0.60	4	2.26	0.1378	0.0061
nab	15	6.8	0.00	1	1.20	0.1292	0.0044
nod	20	7.0	1.08	12	1.86	0.1223	0.0060
pipe	18	7.0	1.30	20	1.79	0.1558	0.0031
rob	17	7.0	1.28	19	1.64	0.1366	0.0030
rod	21	6.8	1.26	18	2.11	0.1486	0.0036
roach	18	7.0	0.30	2	2.09	0.1074	0.0028
top	21	7.0	2.31	204	1.79	0.1421	0.0047
wed	25	7.0	0.30	2	2.37	0.1312	0.0061
wish	13	6.9	2.04	110	2.32	0.1242	0.0058

Note: **Pos. Seg. Freq.** is position segment frequency (a measure of phonotactic probability); **Biphone Freq.** is biphone frequency (a measure of phonotactic probability). **NF** is neighborhood frequency;

APPENDIX E.3-High Phonotactic Probability Nonwords in IPA used in Experiment 3

Stimulus Nonword (IPA)	Pos. Seg. Freq.	Biphone Freq.	Neighborhood Density	Mean log NF
bɛtʃ	0.132	0.0036	18	21.61
tʃæk	0.142	0.0075	21	65.29
tʃæs	0.167	0.0093	16	37.31
tʃʌl	0.122	0.0051	13	6.54
tʃɒm	0.119	0.0067	11	11.64
dek	0.135	0.0035	26	94.85
dɛʃ	0.133	0.0072	13	42.92
dɝs	0.155	0.0035	15	13.07
dul	0.148	0.0032	27	76.85
dut	0.140	0.0043	29	71.48
fætʃ	0.134	0.0049	13	15.23
fob	0.122	0.0075	10	999.90
dʒɪv	0.134	0.0046	11	58.55
kʌʃ	0.140	0.0047	14	67.14
kʌz	0.152	0.0053	12	129.33
koθ	0.149	0.0067	15	71.33
lɛz	0.127	0.0054	13	93.85
mʌv	0.120	0.0051	14	98.50
mʌz	0.117	0.005	15	99.33
mɒtʃ	0.126	0.0066	11	105.09
mɒdʒ	0.128	0.0073	11	8.73
nos	0.152	0.0058	14	259.14
rædʒ	0.140	0.0057	15	17.33
rɛʃ	0.131	0.0086	11	35.36
sʌʃ	0.149	0.0063	14	236.07
ʃɒl	0.144	0.0065	13	39.69
ʃɛm	0.132	0.0057	10	204.30
sɝg	0.145	0.004	8	35.88
sɝp	0.164	0.0042	16	20.75
sɝθ	0.134	0.0039	15	56.20
tæθ	0.131	0.0049	15	16.67
θʌl	0.120	0.005	11	6.36
vɪʃ	0.126	0.0051	5	32.60
vɪθ	0.126	0.005	6	1221.17
zin	0.131	0.0033	15	47.93
zaɪn	0.133	0.005	16	52.13

Note: **Pos. Seg. Freq.** is position segment frequency (a measure of phonotactic probability); **Biphone Freq.** is biphone frequency (a measure of phonotactic probability). **Mean log NF** is mean of neighborhood log frequency

APPENDIX E.4-Low Phonotactic Probability Nonwords in IPA used in Experiment 3

Stimulus Nonword (IPA)	Pos. Seg. Freq.	Biphone Freq.	Neighborhood Density	Mean log NF
biv	0.107	0.0037	17	407.77
tʃɛf	0.102	0.0026	10	36.90
tʃɛg	0.100	0.0026	8	29.00
tʃop	0.095	0.0013	17	22.35
tʃut	0.097	0.0028	17	14.59
dʌdʒ	0.102	0.0029	18	54.67
dɒʃ	0.120	0.0025	12	9.33
dog	0.119	0.002	16	29.31
doθ	0.109	0.0017	15	92.00
dɝm	0.126	0.0031	16	21.63
fʌf	0.106	0.0026	15	9.87
fɛʃ	0.127	0.0029	10	39.70
fɝv	0.095	0.0029	13	44.92
gek	0.109	0.0032	22	99.05
dʒaɪl	0.122	0.003	17	12.71
kɝtʃ	0.125	0.0026	17	35.71
kɝθ	0.125	0.002	16	26.75
lʌz	0.093	0.0028	13	66.54
mʌdʒ	0.107	0.0044	17	62.88
mʌθ	0.104	0.0041	16	78.00
mum	0.129	0.0027	18	46.06
næz	0.123	0.0025	13	768.92
nʌp	0.100	0.0026	14	149.57
rʌz	0.110	0.0036	16	65.13
rɒg	0.129	0.0017	12	12.42
ʃɒb	0.096	0.0026	15	34.67
ʃɛb	0.109	0.0016	8	18.38
saʊʃ	0.120	0.0007	5	49.80
θæʃ	0.094	0.0022	13	5.62
taʊl	0.128	0.0009	19	38.79
vætʃ	0.110	0.004	11	13.09
væʃ	0.110	0.0045	13	8.00
jɝl	0.106	0.0011	12	21.42
zɪtʃ	0.107	0.0018	10	13.90
zɪθ	0.106	0.0016	6	1222.33
zun	0.121	0.0027	14	31.36

Note: **Pos. Seg. Freq.** is position segment frequency (a measure of phonotactic probability); **Biphone Freq.** is biphone frequency (a measure of phonotactic probability). **Mean log NF** is mean of neighborhood log frequency

APPENDIX E.5- Nonword foils in IPA used in Experiment 3

tʃæl
dis
fɝs
kʌv
kɒtʃ
kɒθ
kof
mɛtʃ
mɛz
mɝn
maʊn
pɒf
pɒdʒ
pɒθ
pim
pob
pof
pog
ʃæs
sɛʃ
θæl
θæs
θɒn
vɛd
zæl
zæs
zɪt
dotʃ
fɝg
gɒf
gɛb
gɛg
giz
gop
gaʊl
lɒʃ
nædʒ
næθ
nɝt
ʃiv
θɒb
θɒg
θod
vaʊs
wætʃ
wæθ
wʌk
wʌp
wɒb
wɒf
zæf
zɛb
zɛg
bæb
bɔɪn
baʊn
tʃæz
tʃɛd
tʃɪθ
dɝd
daʊd
fʌtʃ
fɒŋ
fuf
gɒg
gin
gaɪn
dʒʌp
dʒæŋ
dʒuf
kɛp
kaɪs
kim
lɝg
leb
ludʒ
mɪg
mɔɪd
mɝd
næs
nɒf
nik
nɝm
pæʃ
pʌdʒ
pɪb
pɝp
rʌdʒ
riʃ
rɝd
ʃæb
ʃʌd
sæf
sɒʃ
sɛs
sɪdʒ
θæg
θʌp
θɒf
tip
tɪθ
tɔɪd
tup
væp
vɒm
vep
við
vɪdʒ
wʌg
wɛm
wig
wof
jæs
jɒŋ
jib
jum
zæŋ
zɒk
zɛs
zuv

APPENDIX E.6- Nonwords (in IPA) for the DIFFEREENT pairs in Experiment 3

For word stimuli		**For nonword stimuli**		**For foils**		
bæp	bʌθ	bɛm	bið	tʃæv	θɒk	nɒθ
bʌv	bɛdʒ	tʃæm	tʃɛp	dib	θɒp	niʃ
bɒl	bɝl	tʃæg	tʃɛʃ	fɝb	θov	nɝz
bɛs	dæt	tʃʌv	tʃof	kʌn	vaʊl	pæv
dʌt	dʌs	tʃɒʃ	tʃus	kɒŋ	wæp	pʌθ
dɒp	dʌʃ	dev	dʌp	kɒdʒ	wæz	pɪf
dɒm	dɒz	dɛp	dɒŋ	koz	wʌm	pɝθ
dɛg	dob	dɝg	dof	mɛk	wʌʃ	rʌs
dɛdʒ	doð	duð	dod	mɛl	wɒŋ	rit
dɪv	fɛf	dub	dɝp	mɝs	wɒθ	rɝs
fæp	fep	fæʃ	fʌv	maʊt	zæb	ʃæp
fot	gæm	fof	fɛt	pɒb	zɛtʃ	ʃʌz
gʌŋ	gɪk	dʒɪtʃ	fɝp	pɒtʃ	zɛk	sæθ
kæk	kus	kʌg	ges	pɒz	bæf	sɒθ
kog	læl	kʌθ	dʒaɪk	pidʒ	bɔɪs	sɛg
koð	læn	kob	kɝf	pom	baʊs	sɪʃ
lɪʃ	mʌn	lɛb	kɝg	poθ	tʃæʃ	θæf
lɪg	mʌp	mʌb	lʌp	pov	tʃɛn	θʌn
mæb	mɛg	mʌŋ	mʌn	ʃæv	tʃɪv	θɒθ
mæf	nætʃ	mɒn	mʌp	sɛf	dɝn	tiv
mɒg	nɒp	mɒʃ	mudʒ	θæd	daʊθ	tɪd
mɒt	paɪv	nog	næd	θæp	fʌŋ	tɔɪs
pɒn	rɒʃ	ræf	nʌd	θɒm	fɒk	tuð
rɛz	rɒtʃ	rɛm	rʌtʃ	vɛl	fuk	væl
rɪn	rok	sʌg	rɒp	zæʃ	gɒm	vɒt
sæl	tɒθ	ʃɒf	ʃɒŋ	zæv	git	vev
sʌθ	wɛp	ʃɛtʃ	ʃɛn	zɪm	gaɪt	vim
tædʒ	wɪb	sɝm	saʊdʒ	dok	dʒʌs	vɪs
		sɝb	θæv	fɝdʒ	dʒædʒ	wʌz
		sɝz	taʊz	gɒk	dʒuv	wɛv
		tæl	væθ	gɛd	kɛt	wim
		θʌʃ	væg	gɛl	kaɪv	wos
		vɪf	jɝm	gif	kiv	jæl
		vɪk	zɪs	goθ	lɝs	jɒp
		zif	zɪg	gaʊθ	lef	jif
		zaɪt	zuk	lɒm	luθ	jut
				næf	mɪp	zæn
				næl	mɔɪn	zɒθ
				nɝg	mɝm	zɛl
				ʃid	næm	zup

APPENDIX F.1 Analysis of raw reaction times from Experiment 1 with participants as the random variable subjected to a 2 x 2 mixed-design ANOVA with neighborhood density as a within-subjects factor and accent type as a between-subjects factor.

	Mean Reaction Time /ms (s.d.)	
	Neighborhood Density	
Accent Type	Dense	Sparse
Native	974.44 (119.18)	952.43 (104.25)
Accented	1155.43 (142.60)	1116.49 (142.60)

	ANOVAs	
Effects	$F_1(1, 38)$	*p-value*
Neighborhood Density	15.13	< .0001
Accent Type	18.81	< .0001
Neighborhood Density x Accent Type	1.17	0.287

APPENDIX F.2 Analysis of raw reaction times from Experiments 1 and 2 with participants as the random variable subjected to a 2 x 2 x 2 mixed-design ANOVAs with neighborhood density as a within-subjects factor, and accent type and task as between-subjects factors.

	Mean Reaction Time /ms (s.d.)			
	Lexical Decision Task		AX Task	
Accent Type	Dense	Sparse	Dense	Sparse
Native	974.44 (119.18)	952.43 (104.25)	806.46 (109.08)	808.36 (104.93)
Accented	1155.43 (142.60)	1116.49 (142.60)	813.55 (64.06)	821.78 (91.43)

	ANOVAs	
Effects	F_1 (1, 76)	*p-value*
Neighborhood Density	5.46	0.022
Accent Type	13.87	< .0001
Task	93.37	< .0001
Accent Type x Task	10.93	0.001
Neighborhood Density x Task	10.68	0.002

	Bonferroni Test	
Post-hoc Effects for Accent Type x Task	F_1 (1, 76)	*p-value*
AX vs. LD in Native condition	20.21	< .0001
AX vs. LD in Accented condition	84.09	< .0001
Native vs. Accented in LD Task	24.71	< .0001
Neighborhood Density x Task		
AX vs. LD in Sparse condition	76.14	< .0001
AX vs. LD in Dense condition	102.87	< .0001
Dense vs. Sparse in LD task	15.71	< .0001

APPENDIX F.3 Analysis of raw reaction times from Experiment 3 with participants as the random variable subjected to a 2 x 2 x2 mixed-design ANOVAs with phonotactic probability (PP) and lexicality as within-subjects factors, and accent type as a between-subjects factor.

	Mean Reaction Time /ms (s.d.)			
	Words		Nonwords	
Accent Type	High PP	Low PP	High PP	Low PP
Native	762.79 (59.47)	763.58 (74.37)	802.47 (82.3)	781.50 (76.98)
Accented	797.83 (80.60)	772.79 (93.15)	847.62 (100.93)	835.81 (79.54)

	ANOVAs	
Effects	$F_1(1, 38)$	*p-value*
Phonotactic Probability	5.64	0.023
Lexicality	50.89	< .0001
Accent Type x Lexicality	5.34	0.026

	Bonferroni test	
Post-hoc Effects for Accent Type x Lexicality	$F_1(1, 38)$	*p-value*
Words vs. Nonwords in Native condition	11.63	0.002
Words vs. Nonwords in Accented condition	44.61	< .0001

Page left intentionally blank