Table of Contents

List of Tables and Figures

Acknowledgments

I would first like to thank my advisor, Dr. Juhani Järvikivi, for his patience and expertise at all levels of this research project. I would also like to thank Dr. Adriana Hanulíková for providing data files from her research and the text used in the production experiment in this study.

I am very thankful to the faculty and staff in the Department of Language and Communication Studies at NTNU for the generous use of their recording studio. Thank you to Dr. Vim van Dommelen, Dr. Jacque Koreman, and Christian Dillner Hagen for their time and patience.

A special thanks to Soheila Shafiei for her help in creating the experiment and in gently coercing all her Iranian acquaintances to act as participants in the study.

I recognize that this thesis would not have been possible without the generosity of the Norwegian people and government. Thank you Norway, for allowing an American to slip across the border and spend two years in your wonderful country. Tusen takk!

Abstract

Spoken-word recognition of foreign-accented /θ/ by L2 English listeners

A production task and a cross-modal priming experiment were used to investigate the influences of accented pronunciation and linguistic experience on recognition of non-native spoken words. A masked face prime image was also shown to determine if seeing a face that matched the ethnicity and gender of the speaker of the auditory prime would facilitate accented word recognition. Chinese and Iranian learners of English heard words containing either an /s/ or a /t/ substitution for the interdental fricative /θ/. A mixed-effects model analysis of the response time data showed that participant groups differed markedly depending on the prime condition and the speaker. Chinese participants showed significant facilitation for /s/ primes spoken by the Chinese speaker, inhibition for /s/ primes spoken by the Iranian speaker, and no priming effects for the /t/ substitution. The Iranian participants, on the other hand, appear not to have been affected by the substitutions for /θ/ but rather by the accent itself, showing a marginally significant facilitation for both /s/ and /t/ variants when they were spoken by the Iranian speaker. This study provides empirical evidence that participants with different L1s do not necessarily process accented words produced by speakers with different L1s in the same way: segmental substitution may weigh more heavily for some groups, while others may be affected by phonetic detail of the accent as a whole.

Chapter 1 Introduction

"U.S. citizens rows tirty to sixty," shouted the woman in heavily-accented English. I was in the international arrivals area of the Los Angeles airport last year, feeling a bit confused. When the other passengers and I approached the immigration counters, the woman in uniform continued to shout the same announcement. I looked at the numbers above the counters and saw nothing higher than 20, so I assumed she meant rows 13 to 16. I was not expecting to hear accented English at that moment, and without the aid of written numbers above the counters, I may have needed to ask her for clarification. But what of my fellow passengers, I wondered. Did their brains go through the same process of interpretation and inference? Was the woman's accent more understandable for people with the same accent? If their first language (L1) did not contain the voiceless 'th' or /θ/ sound, as I imagine the airport employee's L1 did not, was the process the same, or was it possibly easier for them?

Having worked for many years teaching English as a foreign language, the topic of accents and communication between non-native speakers has long interested me. I have often taught groups of people from various countries, and it becomes readily apparent that students' L1 has a profound effect on their ability to acquire English phonology, especially in adulthood. Students sometimes comment in the beginner and intermediate stages of language learning that they are better able to understand each other than they are native English speakers. This makes sense if we consider that learners of any language who have similar proficiency are likely to use vocabulary and grammatical structures that are also of similar complexity. But what of the phonology? If both speakers' languages do not contain the /θ/ sound as in my example above, and neither speaker is able to produce the /θ/, are words that include the sound stored and processed differently in their brains? In a nutshell, that is what this study sets out to explore.

This question has obvious theoretical implications for understanding how a second language is processed, but as a language teacher, my hope is that it might also lead to some practical insights. Now that approximately 80% of English speakers in the world are non-native speakers (Graddol, 2000), some researchers have begun to question the long-held belief that native-speaker pronunciation norms must be retained in the classroom (Jenkins, 2002). Jenkins (2002) created a list of the English sounds that, when mispronounced, most often caused breakdowns in communication between her students. In her classroom experience, /θ/ was one of the sounds that, when substituted with /s/ or /t/, was less likely to lead to miscommunication. Accented pronunciation of vowels, on the other hand, often led to misunderstandings (Jenkins, 2002).

It was precisely Jenkins' assertion that "the replacement of /θ/ with /s/ was not at all problematic" (2002, p. 88) that led me to the topic of this thesis. While her data were collected over a number of years in the classroom, I saw the need for a controlled experiment where her claims could be put to the test. If /θ/ substitutions are readily comprehensible among non-native speakers, then maybe classroom time is better spent on phonemic contrasts in English that more often lead to miscommunication.

Over the last two decades, there has been a growing body of research into L2 processing, and with /θ/ being a notoriously difficult sound for English learners to acquire, that sound has been the focus of a number of studies (Hanulíková & Weber, 2011; Reis, 2006; Wang, Behne, & Jiang, 2008). In Chapter 3, the reader will find a broad examination of the research underpinning this thesis, but here I will focus on a few of the more recent studies that have informed my own research goals. Let me clarify that, although my interest in this topic stems from my work as a language teacher, the study detailed in this thesis is a psycholinguistic experiment that only has an indirect connection with language teaching and learning research.

First, in their study with German and Dutch L2 speakers of English, Hanulíková and Weber (2011) demonstrated that experience with accented pronunciation appears to facilitate recognition of variants in the L2. When they speak of "experience", they refer to a logical assumption that a German speaker of English is likely more acquainted with German-accented English, through hearing her own accent and that of her compatriots, than she is with other foreign accents. Using eyetracking, they showed that German listeners looked preferentially at English target words that began with /θ/ when they heard the same words pronounced with an /s/ in place of the /θ/, a substitution typical of German-accented English. The same was true for Dutch speakers when they heard a /t/ substituted for the /θ/. Interestingly, this facilitation occurred regardless of whether the /s/ or /t/ variant was produced by a German speaker or a Dutch speaker. The authors also undertook a perception experiment, and determined that /f/ was more easily confused with /θ/ than /s/ or /t/ for both Dutch and German listeners. However, neither group of listeners experienced the same degree of looking preference for /θ/ words that were pronounced with an /f/ substitution. Also revealingly, they did not find a direct correlation between the way each *individual* participant pronounced /θ/ in a production study and the looking preferences for that participant. The participants in their study were living in their home countries, and therefore one can assume that they had more experience with their compatriots' pronunciation of English than with the accent of the other group. Hanulíková and Weber (2011) propose that L2 listeners who are resident in their home countries may have encountered the accented variant, in this case /sɪŋk/, before encountering the canonical form /θɪŋk/. However, the lack of a correlation between participants' production of /θ/

and their processing as revealed in the eyetracking experiment raises an important question about the relationship between production and comprehension, a question I seek to address in this study.

Weber, Broersma and Aoyagi (2011) found similar experiential effects in the processing of accented English by Dutch and Japanese speakers, but unlike Hanulíková and Weber (2011), they also found facilitation for foreign-accented words that were perceptually confusable with the standard pronunciation. This contrast is due to the specific sounds that they investigated. Hanulíková and Weber (2011) looked at /θ/ in relation to /f/, /s/, and /t/, all sounds which exist in both German and Dutch. In contrast, Weber et al. (2011) had Japanese participants listen to Dutch-accented English words, in some of which /æ/ was replaced by /ɛ/. As neither of these two sounds exists in Japanese, it is to be expected that Japanese listeners might experience facilitation for the Dutch-accented variant, as they were likely unable to distinguish the difference between the variant form and the standard form.

Both Hanulíková and Weber (2011) and Weber et al. (2011) suggest that experience with accented English has an effect on recognizing words carrying the same accent. This has been shown to be the case in various L1 studies, and these will be discussed in more detail in the next chapter. I have opted to see the role of experience in terms of the statistical frequency of variant forms and their inclusion in the lexical representation of a given lexical item. Ranbom and Connine (2007) posit that variant forms are stored together with canonical forms, but the strength of the link between them is determined by the frequency of the variant form. Their hypothesis underlies the predicted role of experience in this study. That is, I hypothesize that the participants will show facilitation for the variant forms of /θ/ words that match the accent from their home country precisely because they are likely to have heard that variant pronunciation, or accent in this case, more frequently than an accent from a different country. My own research sets out to see whether similar experiential effects can be demonstrated with participants from different language groups, in this case Persian[1] and Chinese, while also assessing the relationship between an individual's production and processing of variant forms. I will also examine Hanulíková and Weber's (2011) finding that /θ/ substitution variants produced priming effects even when the speaker was essentially mimicking a different accent than her own.

In addition to exploring the linguistic factors that determine processing of accented English, I also investigate a non-linguistic factor of potential significance by using a masked prime image of a face. Face priming has been used extensively in social psychology research to investigate ethnic

1 I have opted to use the term *Persian* rather than *Farsi* to refer to the language of the Iranian participants in this study. *Persian* can be used to refer to mutually intelligible dialects spoken in Afghanistan and other areas of Central Asia, in addition to being the dominant language of Iran, and it appears to be the term preferred by scholars when referring to the language in English (Stilo, Talattof, & Clinton, 2005). I will also use the term *Chinese* to refer to Mandarin Chinese, the official language of the People's Republic of China.

and gender stereotypes, among other topics. For example, Eberhardt, Goff, Purdie and Davies (2004) used masked primes of Black and White faces to investigate racial stereotypes in the United States. They found that in comparison with White face primes, Black faces reduced the amount of time White participants needed to detect crime-relevant objects. They posit that the association in White Americans' minds between Blacks and crime served to create "perceptual processing biases" (Eberhardt et al., 2004, p. 879) that affected the detection of the target objects. If facial images of an ethnic group can activate a concept such as crime, it is possible that ethnically distinct facial primes may also affect how participants process accented spoken language. In this study, I set out to investigate whether such non-linguistic, socio-cultural information affects how accented speech is perceived by L2 listeners.

1.1 Aims of the Research

The main questions of the study are:

1) Does production of the English interdental fricative /θ/ correlate with performance on a lexical decision task that includes accented production of the same phoneme?
2) Does experience with the accent of English from one's home country facilitate word recognition of similarly-accented English?
3) Does it matter if the accent is imitated or authentic?
4) Does a subliminal face that matches the ethnicity and gender of the speaker facilitate accented word recognition in the L2?

Regarding questions (2) and (3), I hypothesize that listeners will recognize /θ/ words more easily, and therefore more quickly and with fewer errors, when they are spoken with a substitution typical of their own accent, regardless of whether the accent is imitated or authentic. This is due to their increased exposure to the accented English common to their home country, either through hearing their own accent or that of their compatriots during their years of English study there. Words spoken with a /θ/ substitution that does not match the listener's, on the other hand, are expected to be more difficult to recognize and therefore elicit slower response times (RTs) and more errors. In response to question (1), while all participants are hypothesized to react in this way, the effect is anticipated to be even clearer among participants who produce more accented tokens of /θ/ themselves. This is because participants who produce the accented variants of /θ/ are more likely to have a stronger link between the variant and canonical forms stored in their mental lexicon when compared with others who were exposed to the variant pronunciation but do not produce it themselves. This topic will be explored in more detail in the literature review in Chapter 2.

The use of the face prime is exploring new territory, and therefore a well-founded hypothesis is difficult to come by. Nevertheless, in line with the face-priming research mentioned above, I anticipate that a subliminal face that matches the accent and gender of the spoken prime word will facilitate recognition and thereby lead to faster RTs.

I have used the terms *speech perception* and *speech processing* in the first part of this introduction. *Speech perception* is often used to describe the biological mechanisms of hearing speech and the processes involved in segmenting the continuous stream of speech into units of some kind, enabling access to the mental lexicon. The term *speech processing* is often used in the broadest sense to include *speech perception* as well as syntactic parsing, referring to the entire process of hearing and understanding continuous speech (Fernandez & Cairns, 2011). The cross-modal priming experiment in this study deals with word recognition; there is no contextual information and no syntactic parsing involved. However, I understand the steps involved in segmenting and mapping spoken input onto representations in the mental lexicon as a process (Norris, 1994), and will therefore use the term *processing* in this thesis.

This study included a production experiment and a cross-modal priming lexical decision experiment with both a visual face and an auditory word prime in each trial. The participants were 12 Iranian and 12 Chinese volunteers. The production experiment involved the participants reading a short English text with 25 occurrences of /θ/ words. Their speech was recorded and analyzed to determine what substitution, if any, they use when producing words with /θ/. For the cross-modal priming experiment, the participants were sequentially assigned to one of six versions of the experiment in a three (prime type: Iranian, Chinese, or control) by two (accent of prime: Persian or Chinese) repeated measures factorial design, with response times (RTs) to the visual target word as the primary dependent variable. Twenty-four /θ/ target words were used, 12 with the /θ/ in a word-initial position and 12 in a word-final position. The prime type for the Persian accent involved substituting a /t/ for the /θ/ in the 24 experimental items which were chosen based on frequency, and the fact that the substitution did not result in a different English word but rather in a pseudoword. The Chinese accent used an /s/ in place of the /θ/. The common substitution of /t/ by Persian speakers and /s/ by Chinese speakers has been documented (Chang, 2001; Wilson & Wilson, 2001). It is important to note that while neither Chinese nor Persian has the /θ/ phoneme, both languages have the /s/ and the /t/ phoneme.

This study seeks to contribute to our understanding of the roles of experience and production in the processing of L2 English words, as well as to investigate the potentially promising influence of non-linguistic factors such as a still facial image. A number of studies have looked at /θ/ substitutions in L2 English (Hancin-Bhatt, 1994; Hanulíková & Weber, 2011; Reis, 2006), and this research seeks to extend Hanulíková and Weber's (2011) research by looking at participants from two language groups, Persian and Chinese, which are more distant from each other than the German and Dutch languages represented in their study. The production experiment in this research is modeled on the experiment by Hanulíková and Weber (2011), while the cross-modal priming experiment is modeled on a similar experiment by Weber et al. (2011) detailed above. The field of non-native speech processing by L2 listeners is relatively new, and there is a need to test the results of previous research with different language groups to determine whether their findings are language specific or can be generalized to all languages. This research seeks to do that.

Regarding the use of facial primes in this project it is my hope that this will open the door to an interesting possibility in psycholinguistic research. Of course, there has been extensive research into the role that seeing a moving mouth and face has on speech processing (Chen & Massaro, 2004; Massaro, 1998; Navarra & Soto-Faraco, 2007; Soto-Faraco et al., 2007). Some of this research has demonstrated that seeing a mouth produce difficult sounds in the L2 can help learners to improve their pronunciation of those sounds (Massaro, 2003). However, to the best of my knowledge, there has not been research conducted on the effect of a still face on L2 language processing, specifically on the possibility that a face from a particular ethnic group may in some way facilitate comprehension by activating the experience that the listener has with the corresponding foreign accent. The results of this study may therefore potentially increase our understanding of the interplay between linguistic and visual experience.

The recordings from the production study were analyzed by me and two other native English speakers. I made perceptual judgments on the participants' production of /θ/, categorizing them as /s/,/t/, or other, and some of those results were compared with the judgments of the other raters. The results of the cross-modal priming experiment were analyzed using linear mixed-effects models. Mixed-effects models have recently been shown to offer advantages in the analysis of repeated measures data in psycholinguistic research (Baayen, Davidson, & Bates, 2008), and they will be discussed in more detail below.

1.2 Organization of the Study

This thesis consists of five chapters. After this introduction, Chapter 2 is a review of general research on the topic of spoken-language processing, as well as a more specific look at the L2 processing research that underpins this thesis and informs the research questions contained herein. Chapter 3 describes the methods used in the study for data collection and analysis, as well as a justification of these choices. In Chapter 4, the results from the two experiments are presented and discussed as well the findings of the subsequent data analysis. Chapter 5 includes a summary of the findings, the implications for research in this area, the limitations of the study, and suggestions for future research.

Chapter 2 Literature Review

This study explores processing of non-native spoken English by non-native listeners. Recent research in this area is based on a long history of previous research into L1 processing. Therefore, this literature review will explore the basics of what we know about spoken-language processing, followed by a look at more recent research into L2 processing.

2.1 The Basics of Spoken-language Processing

The problem of how humans decode the nearly continuous stream of spoken language is certainly a complex one. Some aspects of this process are agreed upon by researchers, while other areas are still the topic of heated debates. One of the fundamental pieces of the puzzle upon which most researchers agree is the idea that spoken language simultaneously activates competing word candidates, and that this process of activation and competition is the bridge between the speech signal and lexical access (Marslen-Wilson & Warren, 1994). While most researchers agree that activation and competition are fundamental in speech processing, they most certainly do not agree on exactly how this process takes place.

The concept of multiple lexical activation arose from Morton's (1969) logogen model and Quillian's (1969) attempt to model semantic memory on a computer. In Morton's model, a logogen is a location in the brain associated with a specific word or concept that contains orthographic, semantic, and phonological information. Morton proposed that both spoken and written input feed to a single set of logogens, and when these pass a certain threshold, the word or concept becomes available for use. Quillian, on the other hand, was concerned predominantly with semantic memory, and his goal was to create a computer program that could comprehend written text in both literal and figurative usage, and to use that program to gain insight into how humans process language (1969). In Quillian's model, a concept is seen as a node in a network, with links to other nodes that vary in strength. For example, the concept of a *typewriter* would have a strong link to the concept of *machine*, while the concept of *machine* would likely have a weaker link to signal that a typewriter is one type of machine. Quillian coined the term "activation tag" (1969, p. 463) to describe the manner by which semantic links could be traced to comprehend a written text. As each word is analyzed, it activates all possible meanings and concepts associated with that word, and this activation spreads like a "fan" (Quillian, 1969, p. 464). When an intersection of nodes is reached, signaling shared meaning, it is traced back to the original word and evaluated for goodness of fit based on syntax and context.

Throughout the early 1970s, numerous researchers designed experiments to empirically test (and attempt to refute in many cases), Quillian's model (1969). Influential proponents of the model were to be found in Collins and Loftus (1975), whose paper on the "spreading-activation theory" of semantic processing brought the concept to the forefront in linguistics.

Much of this research into activation dealt with semantic processing. One of the early attempts to examine how activation might apply to phonological processing, which is arguably a crucial hurdle that must be passed *before* semantic processing can take place, was the cohort model (Marslen-Wilson & Tyler, 1980). The basic concept is, when listeners hear the /k/ of /kæt /, their brains will "activate" not only the word *cat* but also all other words that begin with the same sound such as *coffee* and *catapult.* These competing hypotheses are considered by the listener until additional information both rules out the wrong ones and more strongly activates the likelier ones (McQueen, 2007). Later research showed that this process is complicated by the frequency and number of similar-sounding words (Luce & Pisoni, 1998). In my example above, the /k/ in /kæt / would activate many hundreds if not thousands of potential word candidates in English because this is a very common word beginning, whereas the sound /z/ in *xylophone* would activate far fewer candidates. This model posited that word onsets are crucial in lexical access, and this was demonstrated in various experiments (e.g. Marslen-Wilson & Zwitserlood, 1989). This makes logical sense, as the speech stream is sequential; many long words reach their uniqueness point before the end of the word, making a substitution at the end less important. However, later versions of the cohort model allowed for constant re-analysis of the activated word candidates, thereby permitting that words that were mispronounced at the onset could still be recognized (Gaskell & Marslen-Wilson, 1997).

While researchers in the 1970s and early 1980s were investigating semantic and phonological processing as discussed above, there were also those who decided to look at the relationship between orthography and spoken-language processing. Seidenberg and Tanenhaus (1979) pointed out that both the logogen model and the spreading-activation model of processing included an orthographic representation of a word that is accessed in the same way as the semantic and phonological representations. Therefore, they set out to test what they called their "counterintuitive prediction" (Seidenberg & Tanenhaus, 1979, p. 547) that the orthographic code would be activated in auditory word recognition. They found that indeed, even when primes and targets were both auditory, orthographic differences affected response times. For example, when participants heard primes such as *toast* and *ghost* that were orthographically similar or dissimilar to the target *roast*, they showed faster RTs for the similar primes, in this case *toast (Seidenberg & Tanenhaus, 1979)*. These findings are bolstered by research conducted with illiterate adults (Morais, Cary, Alegria, & Bertelson, 1979). Morais and colleagues found that in contrast to literate adults living in a similar environment, illiterate adults were unable to delete or add phonemes from pseudowords, leading them to propose that "awareness of speech as a sequence of phones is thus not attained spontaneously in the course of general cognitive growth" (Morais et al., 1979, p. 323). In other words, learning to read profoundly affects the way we process not only written language, but spoken language as well.

Other researchers continued looking for clues to understand the overall picture of language processing, and with the growth of interest in neural networks in the 1980s, a connectionist model of language processing called TRACE was developed by McClelland and Elman (1986). One of the limitations of previous models was the reliance on the idea that the speech stream should be broken into discrete phonemic units in order to enable lexical access (Dahan & Magnuson, 2006). However, the existence of such phonemes was complicated by previous research showing that there is parallel transmission of information in adjacent phonemic units (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). In other words, phonemes, if they exist, vary depending on the sounds that surround them. The TRACE model addressed this issue by positing sub-phonemic units called features that are activated and compete with each other so that eventually the phoneme that carries the most positive evidence wins the competition (Dahan & Magnuson, 2006). These features are envisioned as nodes in a network, and the activation can be both positive and negative. Activation spreads upward from features that match the input, and are thus positively activated, to the phonemic level and thence to the lexical level. The negative activation, or inhibition, does not occur at the feature level but rather once potential candidate words have been activated. The lexical candidates with stronger activation, due to a closer match with the input, will inhibit other potential lexical candidates that do not match the input as well (McClelland & Elman, 1986). Therefore in TRACE, unlike in the cohort model, the sequential nature of speech is seen as less crucial. For example, the input /bləzənt/ will activate the word *pleasant* despite the mismatch on the first phoneme because there is a lot of overlap between the two, and because there is no existing word *bleasant (McClelland & Elman, 1986)*. An eye-tracking study by Allopena, Magnuson and Tanenhaus (1998) supported the predictions of the TRACE model by establishing the time course of activation for competing words with phonetic similarities. Although this concept of "nodes in a network" harkens back, at least in my mind, to Quillian's (1969) model of semantic processing, McClelland and Elman (1986) do not cite Quillian's research in their original description of the TRACE model.

TRACE also differs from previous models in that it allows feedback from the lexical to the phonemic level. To use an item from my own experiment, suppose a listener were to hear the sound /hɛl/ followed by an ambiguous sound between a /s/ and a /θ/. Although both phonemes may be activated at first based on the bottom-up information in the signal, the fact that /hɛlθ/ is a word while /hɛls/ is not would send information back to the phonemic level of processing to favor /θ/ over /s/.

The proposal that information from the lexical level feeds back to a pre-lexical level was met with an article by dissenting researchers which included the subtitle "feedback is never necessary" (Norris, McQueen, & Cutler, 2000). They argued that feedback from the lexical level into the phonemic and pre-lexical levels could lead to corruption of the original signal. Their proposed model, Merge, has a "bottom-up priority rule" (Norris et al., 2000, p. 312); it does not completely do away with feedback, but rather posits a separate, parallel phonemic level. Importantly, this parallel set of phoneme nodes can integrate information from above and below without altering the pre-lexical information from the speech stream (Dahan & Magnuson, 2006).

Recent research by Vitevitch and Luce (1999) has provided a more nuanced view of the way that activated candidates compete among each other, as well as a clearer picture of the difference between pre-lexical and lexical processing as described in some of the models above. The authors set out to resolve a contradiction. The Neighborhood Activation Model (Luce & Pisoni, 1998) proposed that words with many similar sounding neighbors would be recognized more *slowly* due to increased competition among potential word candidates. In contrast, earlier research by some of the same authors (Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997) had shown that phonotactic patterns that occur more frequently lead to *faster* processing of nonwords, but *slower* processing of real words. Vitevitch and Luce (1999) hypothesized that this contradiction arises from differences in processing at the pre-lexical and lexical levels. For real words, increased lexical competition overshadows any benefits that the corresponding higher phonotactic probabilities would provide to speed up processing. In contrast, nonwords experience much less lexical competition and therefore phonotactics appear to dominate in their processing. In addition, they assert that only the Shortlist (Norris, 1994) model of processing contains the requisite architecture to account for their findings: "opposite effects of probability and density as a function of lexicality" (Vitevitch & Luce, 1999, p. 401).

The Shortlist model (Norris, 1994), like the Merge model mentioned above, was proposed in response to what were seen as shortcomings of the most widely-applied model in the 1990s, TRACE (McClelland & Elman, 1986). Like TRACE, the Shortlist model relies on competition between lexical candidates, but it differs in not allowing feedback: Shortlist is "entirely bottom up in its operation" (Norris, 1994, p. 190).

In recent years, Shortlist has undergone a major revision. Shortlist B is similar to the original in that it is bottom-up, does not include feedback, assumes competition among multiple lexical candidates, and includes phonologically abstract pre-lexical and lexical representations (Norris & McQueen, 2008). However, Shortlist B differs in no longer assuming discrete phonemic units as input, and most importantly, it replaces the long-used concept of "activation" with Bayesian computations involving likelihood and probability. The model relies on prior probabilities of variable pronunciation, frequency, and context to perform lexical access and segmentation in an optimal Bayesian decision-making process. The authors concede, however, that the model still needs to be tested empirically.

A fourth theory of speech perception, which was proposed in 1967 but has recently been reevaluated in light of new research, is the motor theory. This theory was proposed by Liberman et al. (1967), and it differed markedly from others at the time. It postulates that speech perception is closely linked to speech production: the motor system that sends signals from the brain to the articulators in order to speak is also recruited when we listen to speech. In their recent review of the motor theory of speech perception, Galantucci et al. (2006) claim that experimental evidence supports the theory in its general sense, that the motor system is accessed in perception.

If one subscribes to the belief that perception is closely linked to production, a natural question to ask is whether the relationship exists in the opposite direction: does production depend in some way on perception? Bradlow et al. (1997) have demonstrated that improvement in perception of L2 phoneme contrasts that do not exist in one's L1 can lead to improvements in production of those same L2 phonemes, but this finding does not necessarily give support to a motor theory explanation of the connection. The motor theory of speech has returned to the headlines in recent years with advances in brain science (Galantucci et al., 2006), but the relationship between perception and production is still much debated. This topic will be revisited below.

2.2 Lexical Segmentation

Words rarely occur on their own in citation form (except in psycholinguistics experiments that is), and this is where the listener's task is further complicated: how to determine where one word ends and the next begins in continuous speech. This topic goes somewhat beyond the scope of this study, but I will give a brief summary.

Both segmental and suprasegmental information come into play in the segmentation of lexical items from the speech stream. The language in question also determines how segmentation takes place: stress-timed languages such as English often have content words that begin with a stressed syllable, a feature that allows the listener to determine where the previous word ends and the next one begins (Cutler & Norris, 1988). Other languages such as Spanish are syllable-timed, while Japanese is mora-timed, and speakers of these languages use syllable and mora information respectively to gauge word boundaries (Fernandez & Cairns, 2011).

Other types of prosodic information have been found to be important in lexical segmentation. Salverda, Dahan and McQueen (2003) used eye tracking to monitor activation of monosyllabic words that had been embedded in larger words (e.g. *ham* in *hamster*). They found that participants' eye movements were affected by the minor differences in vowel length of the monosyllabic *ham* when it was embedded in *hamster* and compared with a normal token of *hamster.* This research demonstrates that fine-grained subphonemic information is also recruited to help listeners segment continuous speech.

For literate people, it is sometimes hard to grasp the fact spoken language is not neatly divided into separate words in the way that written text is, and that *hamster* includes the word *ham.* This may of course differ for speakers of languages such as Japanese that do not put spaces between written words. Nevertheless, this leads to the realization that the models of spoken language processing discussed above become substantially more complex if we consider activation and competition not only of the spoken words in an utterance, but also of short words embedded in longer ones and potential words that bridge the gaps between words. It was precisely this point, that the cohort model ascribed such importance to word onsets without describing how the listener could *find* the word onset in continuous speech, which led Tabossi and colleagues to examine the topic of segmentation (1995). In a cross-modal priming experiment, they used trisyllabic prime words (e.g. *visite,* Italian for *visits*) that contained other bisyllabic words, in this case *visi*, Italian for *faces.* The bisyllabic words were then placed in sentences where they were followed by a word, *tediati*, Italian for *bored,* that began with the third syllable of the trisyllabic word (*te*), thus creating a segmentally ambiguous context. They found that even though the first syllable of the following word, *te̲diati*, showed a different stress pattern than it would have in the trisyllabic word *visite*, and even though *visite* did not fit with the semantic or syntactic context of the sentence, it still appeared to have been activated in the sentence about bored faces. This they gauged by response times to a semantically related target (PARENTI, Italian for *relatives)* presented precisely at the offset of the third syllable. In addition to the insight into processing, this type of research points to a serious shortcoming in many psycholinguistic studies, including my own: the over-reliance on the presentation of single words when this is very rare in natural speech situations.

Be that as it may, I will return to the topic at hand. The above research on the ways in which small changes in properties of phonemes affect lexical access is a natural stepping stone to narrow the focus of this review and discuss some of the other ways in which the speech stream varies.

2.3 Variation in the Speech Stream

The research into how variation is dealt with in spoken-language processing can be broadly divided into accounts that emphasize representations and others that emphasize processing. One of the early representation-based approaches to this topic was proposed by Lahiri and Marslen-Wilson (1991). They hypothesized that abstract lexical entries are "underspecified", in that they only include distinctive information from the surface phonetic realization of the word. For example, the glottal stop /ʔ/ is an allophone of /t/ in many varieties of English, creating the variant form /hɒʔ / in place of /hɒt/. However, the glottal stop is not a phoneme of English and does not create minimal pairs with different meaning. Therefore, according to the underspecification approach, the variant pronunciation /hɒʔ/ of *hot* would not be stored as a lexical entry, but rather only /hɒt/.

Additional representation-based accounts vary in the degree of detail that they attribute to lexical representations. In Goldinger's (1996) episodic model, all the instances of a word that were ever heard are recorded in detail as part of the representation. In contrast, Ranbom and Connine (2007) posit that statistical information comes into play: variant forms are stored together with canonical forms, but the strength of the link between them is determined by the frequency of the variant form. Their hypothesis underlies the predicted role of experience in my own study. That is, I hypothesize that the participants in this study will show facilitation for the variant forms of /θ/ words that match the accent from their home country precisely because they are likely to have heard that variant pronunciation, or accent in this case, more frequently than an accent from a different country.

Processing-based accounts, in contrast, hold that segmental context is used by listeners to deal with variation in spoken language. Most often these accounts look at variation that results from coarticulation. Take the English word *bag* for example. The vowel /æ/ both affects the pronunciation of the neighboring consonants /b/ and /g/, and the consonants in turn affect how the vowel is pronounced (Fernandez & Cairns, 2011). This is known as parallel transmission, and it means that although we like to consider an English vowel sound such as /æ/, for example, as a discrete and stable unit, it is actually a slightly different sound in each and every word in which it occurs depending on the sounds surrounding it.

This applies to consonants as well, so that the /p/ at the end of the word *wrap* is not the same sound as the /p/ at the beginning of the word *paper* (McQueen, 2005). There is additional variation in the realization of each sound depending on its neighboring sounds, and phonological processes such as neutralization, epenthesis, mutation, and assimilation are very often in evidence (Gow, 2001) in running speech. For example, the /n/ sound in the phrase "in Portugal" is different from the /n/ in the word *in* when spoken alone. The /n/ before the /p/ in Portugal is most often pronounced nearer to an /m/ in anticipation of the following bilabial consonant, and this is an example of nasal assimilation. The question that many researchers have sought to answer in recent years is then, what does this variation tell us about how our mental lexicon is organized. Are the canonical form /ɪn/ and variant forms such as /ɪm/ stored separately, or is there a pre-lexical process of smoothing out this variation before we access the mental lexicon?

An important piece of evidence in favor of the "smoothing out" idea is research that demonstrates the categorical nature of our perception of speech sounds. In conjunction with their description of parallel transmission, Liberman et al. (1967) also showed that perception of consonants tends to be categorical, while that of vowels is more continuous. This makes sense if vowel length, which can vary along a continuum, carries information about the surrounding consonants, whereas consonant contrasts in English are often binary between voiced and voiceless pairs.

With this amount of variation, it is amazing that lexical access is even possible. But most variation is not random but rather rule-governed. Indeed, Gaskell and Marslen-Wilson (1996) found that when a sound mismatch occurs in a position where assimilation is permitted in English, normal lexical access can take place. However, if the variation exists in a position where English phonology would not normally license assimilation, access is disrupted. For example, they included the prime word *lean* embedded in a sentence. In the experimental item, *lean* was followed by the word *bacon*, and in the control condition it was followed by *gammon*. As in my example above, nasal assimilation would lead *lean* to be pronounced as /lɪːm/ when followed by *bacon* but not when followed by *gammon.* They observed slower response times for the target [LEAN] when primed with /lɪːm/ followed by *gammon*, the unlicensed context, compared with /lɪːm/ followed by *bacon*, the viable context for assimilated change of /n/ to /m/. They posit that this demonstrates that "the lexical access process is intolerant of small deviations" (Gaskell & Marslen-Wilson, 1996, p. 153). In other words, even though the pronunciation varied by only one feature, access was disrupted when it occurred in an unviable context.

In an effort to gauge the effects of deviant pronunciation on lexical access, and in particular to test earlier claims of Marslen-Wilson and Tyler's cohort model (1980) on the importance of word onsets, Connine, Blasko and Titone (1993) conducted six cross-modal priming experiments. They demonstrated that even when phonemes were altered in word onsets, as long as the substitution only differed from the original phoneme by one or two features (manner, place, or voicing), priming effects were still evident. Although they set out to show that word onsets are not crucial, and they assert in the abstract of the article that auditory word recognition as they understand it "affords no particular status to word-initial phonemes" (1993, p. 193), they concede later that "initial phonemes *may* have a privileged status of sorts"(emphasis in original) (1993, p. 199). They allow that phonemes in word onsets "may require a greater degree of overlap" (Connine et al., 1993, p. 199) than those found in other parts of words. This topic is relevant for my own research, as the experimental items differ in having /θ/ in word initial or word final position, as well as the fact that the /s/ and /t/ substitutions differ in one and two features, respectively, from the standard pronunciation of /θ/.

Before moving on to discuss the treatment of accents in the literature, it is important to step back and notice the disconnect between much of the research into how variation is dealt with in spoken-language processing, and the reality of a face-to-face interaction between a speaker and a listener. In many ways, the language in the experiments detailed in some of the above studies approximates spoken language on the telephone, completely devoid of visual and pragmatic cues. Anyone who has tried to use a foreign language on the telephone can attest to increased difficulty in comprehension when one can not see the speaker's facial expressions nor any visual contextual information. Psycholinguistic experiments such as my own, where one phoneme is manipulated to gauge its effect on processing, or the extensive number of experiments on processing in noise, are valuable in that they seek to push the system of human perception in order to tease apart how it functions. However, it may very well be that processing of isolated words that one hears via headphones is not simply one cog in the larger machine of general processing that occurs in a normal face-to-face interaction, but rather it may be a completely different animal. It is clear that visual cues, when available, are utilized by the brain and help us to process spoken language more accurately (Massaro, 1998; Soto-Faraco et al., 2007). But how is visual information recruited to help the listener deal with non-standard variation in the input? Kraljic, Samuel & Brennan (2008a) set out to test the hypothesis that "speech-perception processes recover invariants not about the signal, but rather about the source that produced the signal" (p. 332). They demonstrated that when only audio was available, and listeners heard non-standard pronunciation of English /s/ during the first half of the experiment, they attributed this variation to idiosyncratic characteristics of the speaker and in essence "learned" to comprehend the non-standard pronunciation. However, when video information was also available, and the same non-standard pronunciation could be attributed to the fact that the speaker had a pen in her mouth while speaking, the participants did not re-shape their perception; it was obvious that the strange pronunciation was due to the pen rather than a stable characteristic of the speaker. The audio-only trials showed a clear "first-impression" bias, and this is relevant for my own research: the participants in this study hear variant pronunciation of /θ/ throughout the experiment and do not receive any additional information that could override the first impression that this is a characteristic of the speakers' speech.

This research on perceptual learning in the L1 is a natural stepping stone leading to research into accents, both foreign and domestic. When a native listener hears an example of assimilation in an unlicensed context, such as *leam gammon* in the above study by Gaskell & Marslen-Wilson (1996), this appears to disrupt processing. But a non-native speaker may produce a similar non-standard pronunciation over and over due to transfer from the speaker's L1, as in /hɛls/ for *health* in this study. If the listener does not see a pen in the speaker's mouth, or some other reason to account for the non-standard pronunciation, Kraljic and colleagues (2008) have demonstrated that listeners will then "re-tune" their perception. Similar retuning and perceptual learning has been demonstrated for manipulated speech in the L1 (Norris, McQueen, & Cutler, 2003), essentially mimicking an accent.

2.4 Accent in L1 and L2

But what determines if a second-language learner will speak with an accent, and how are accents dealt with by listeners? Research by Flege (1999) with Korean immigrants in the United States found that their age of arrival in the country correlated positively with their degree of foreign accent as determined by native listeners. A related study on vowel production among Spanish-speaking immigrants in the United States showed similar results: "accuracy in producing English vowels is related inversely to their age of first extensive exposure to native-produced English" (Flege, 1992, p. 575). These data fit with the long-held belief in language-acquisition research that there is a "critical period" for L1 acquisition that also applies to L2 learning (Johnson & Newport, 1989). In other words, the older you are when you begin to learn a second language, the more likely it is you will speak with an accent.

When a native listener hears accented language, this is an additional example of "variation" in the speech input as discussed above. However regional dialects and foreign accents have been found to differ: while the variation between dialects is often seen in the pronunciation of vowels, foreign accents normally affect all phonemes in the language (Floccia, Butler, Goslin, & Ellis, 2009).

Experimental results have shown a contradictory picture of the way that accents are dealt with by listeners. For example, Bradlow and Bent (2008) demonstrated rapid adaptation to foreign-accented speech by native listeners. Native-English-speaking participants listened to foreign accented English sentences, and they transcribed them without the possibility of replaying them. Over the course of the experiment, not only did the listeners improve their ability to comprehend the Chinese-accented English, in effect "tuning" to the variant pronunciation, but this perceptual adaptation carried over when participants listened to other speakers with the same accent.

Norris, McQueen, and Cutler (2003) also showed rapid perceptual learning, albeit with a digitized sound continuum to create ambiguous phonemes on individual words, essentially mimicking an accent. Crucially, their results showed a difference between effects for words and nonwords. While participants showed rapid adaptation to ambiguous pronunciation of word-final phonemes in real words, similar adaptation was not evident for nonwords. Based on these results, they hypothesize that listeners experience an initial disruption upon hearing an unfamiliar foreign or regional accent. This disruption is due to problems with pre-lexical processing. Once some of the accented speech is recognized, lexical knowledge begins to provide feedback to the pre-lexical level, essentially "retuning" the phonemic categories to match the input. They predict that, after an initial delay, listeners will experience improved comprehension through adaptation. This is in line with the Merge model of spoken-language processing proposed by the same authors (Norris et al., 2000) where feedback is possible, but only along a "separate feedback path for the training signal" (Norris et al., 2003, p. 233) in order to avoid corruption of the original input.

A recent study by Floccia et al. (2009) has questioned this account of rapid adaptation to accented language as proposed by Norris et al. (2003). Floccia and colleagues suggest that listeners experience a delay in comprehension when they are first exposed to accented speech, but their experimental results suggest that they do not subsequently adapt. However, unlike previous studies (e.g. Bradlow & Bent, 2008) showing adaptation where at most two foreign accents were used in the stimuli, Floccia et al. (2009) exposed participants to one foreign and two regional dialects spoken by various speakers over the course of 60 trials in one experiment. I surmise that they may have found the upper limit for adaptation.

Another study on regional dialects has also shown limited flexibility in recognizing non-standard pronunciation. Sumner and Samuel (2009) demonstrated that participants who had experience with the regional accent of English from New York City showed facilitation for recognition of words carrying that accent in a priming experiment. Revealingly, participants who did not have experience with the New York accent did not show priming facilitation for the accented primes, even though they only differed in the r-coloring of one phoneme: NYC prime /brʌðə/ for General American /brʌðɚ/. Participants who lived in the area but did not produce the accented variants themselves did show facilitation. The authors suggest that their results show "a dissociation between production and representation" (Sumner & Samuel, 2009, p. 499). They assert that participants who consistently speak with the r-drop of the NYC accent store representations of both the accented and the standard pronunciation, while other participants who have experience with the dialect but do not produce the variant forms showed priming facilitation, even though the authors assume that they only store a representation of the standard pronunciation.

2.5 Perception and Production in L2

These results point to the enduring question of the relationship between perception and production in L2. It is a common observation in language teaching that production often lags behind perception and comprehension (Sumner & Samuel, 2009). However, research in this area has been complicated by the difficulty in teasing apart the contribution of motoric vs. perceptual factors in L2 production difficulties (Flege, 1992).

Flege has hypothesized that "accurate phonetic perception is a necessary but not sufficient condition for accurate L2 segmental production" (1992, p. 569). The most obvious case of both production and perception difficulty in L2 occurs when neither of two sounds that are contrastive in the L2 occurs in the speaker's L1. A common example of this case involves problems that Japanese speakers have with English /r/ and /l/, both of which map poorly to the apical postalveolar flap /ɽ/ in their L1. Japanese speakers have been documented to have difficulty with both perception and production of these sounds, although with time and training they have also been shown to improve in both areas (Flege, Takagi, & Mann, 1995). There is, however, contradictory evidence showing that "new" vowel sounds in the L2 that do not exist in the L1, in this case for English learners of French, are easier for speakers to mimic than those vowel sounds that are close to vowels in the L1 (Flege, 1987).

A second problem arises for the listener when a sound in the L2 has a counterpart in the L1, but the L2 sound occurs in a novel and unfamiliar context (Flege, 1992). For example, Mandarin Chinese has a contrast between the voiced and voiceless stops /t/ and /d/, but this only occurs in word-initial position (Flege, McCutcheon, & Smith, 1987). If a Chinese speaker were to participate in a perception experiment, she would likely be able to discriminate between the individual phonemes /t/ and /d/. However, this perceptual skill would not translate into the ability to produce the sounds in an unfamiliar location in an English word. A similar situation has been documented with Dutch speakers of English; they show accurate perception of contrasts in unfamiliar positions but difficulty with production (Broersma, 2005).

In general, when the L2 phoneme categories differ from those in the L1, comprehension is normally hindered, especially when there is a distinction in the L2 that maps to a single category in the speaker's L1 (Weber et al., 2011). However, when two phonemes in the L2 correspond with two separate sounds in the L1, comprehension is usually good (Flege, 1993). In either case, however, the amount of competition in L2 listening when compared with L1 listening is normally much higher, and this means slower word recognition (Broersma & Cutler, 2008; Norris, McQueen, & Cutler, 1995). This topic will be discussed in more detail below.

The issue of how perception and production are related cross-linguistically is still an open question. However, it is clear that the L1 has an important and lasting effect on second-language learners' production ability, and this is certainly the case for the Chinese and Iranian participants in this study.

2.6 Listening in the L2

If native listeners can in many cases "retune" in order to accommodate foreign accents and regional dialects, what of non-native listeners? Why is it such hard work to understand a foreign language when one has limited proficiency? One reason is thought to be inaccurate perception of phoneme contrasts in the L2, and the resulting interference of L1 phonology. An influential model that sets out to explain the relationship between a listener's L1 phonology and L2 processing difficulties is the Perceptual Assimilation Model (Best, 1995). This model predicts three different levels of discrimination for non-native phonemes: (1) where a contrast pair in the L2 maps to a similar contrast pair in the L1, discrimination will be very good; (2) where both members of a contrast pair in the L2 map to one sound in the L1, with one member seen as a good match and the other a poor match, discrimination will be worse than in (1) but still good; and (3) where both members of a contrast pair in the L2 map equally poorly to a single sound in the L1, discrimination will be very bad, as in the example of English /r/ and /l/ for Japanese listeners described above. This model is supported by some recent empirical research (Best, McRoberts, & Goodell, 2001; Brannen, 2011).

If we set aside perceptual difficulties, at first glance it might seem that L2 listening should be easier than L1 listening: if one has a limited vocabulary in the L2, then there should be much less competition among potential word candidates. Unfortunately, this does not appear to be the case. In eye-tracking experiments with Dutch learners of English, Weber and Cutler (2004) demonstrated that there is increased lexical competition for non-native listeners when compared with native listeners. They attribute this increase to simultaneous activation of vocabulary from the listener's L1 and inaccurate perception of L2 phonemes, as discussed above, leading to even more competition from spurious candidate words. Specifically, their experiments examined confusable English vowel pairs, and they revealed that the effects of increased competition, at least in this case, are unidirectional. Dutch listeners often confuse English /æ/ and /ɛ/; Dutch has the phoneme /ɛ/, which is a reasonably good match for the English vowel, but it does not have /æ/. They found that spoken input with either vowel (e.g. [pæ] and [pɛ]) activated words beginning with [pɛ], but the reverse was not observed as frequently. They explain this in the following way: "it is as if the phoneme category of the second language that is perceived as nearest to the native category captures all identification responses, while the second-language phoneme that is perceived as further from any native category is simply ignored" (Weber & Cutler, 2004, p. 21).

In a follow up study, Broersma and Cutler (2008) demonstrate that spurious activation in the L2 due to inaccurate phoneme perception also occurs across word boundaries, thus exponentially increasing the amount of competition that can disrupt L2 listening. They examined English word pairs ending in voiced and voiceless sounds. This contrast occurs in English, but in Dutch it only occurs in word-initial position. While Dutch listeners have been shown to perceive this difference when it occurs in word-final position in nonword pairs (Broersma, 2005), Broersma and Cutler (2008) found that listeners experienced "phantom activation" when near words were embedded in or across other words. For example, they excised the near word *groofs* from the fragment *big roofs*, and found that priming with *groofs* activated the English word target GROOVE for Dutch speakers. They point out that when speaking with a native speaker of English, it is unlikely that a listener will encounter isolated occurrences of *groofs*; however, the possibility increases when we consider overlap across words such as *big roofs*, as this situation is very common in running speech. In other words, L2 listening is a monumental challenge.

This cross-linguistic competition has also been demonstrated in the opposite direction, making it clear that it is not possible to partition one's mental lexicon into separate sections for each language. Even when a listener thinks she is completing a monolingual task in her L1, it appears that there is parallel access to lexical candidates in the L2. Bijeljac-Babic, Biardeau and Grainger (1997) demonstrated this using masked orthographic priming. French learners of English saw masked English prime words followed by French targets. When the primes were orthographically related to the targets (e.g. soil - SOIF), participants were slower to respond to the lexical decision task than when the prime was unrelated (e.g. gray - SOIF). This cross-linguistic priming was a function of the participants' proficiency in the L2, with highly-proficient participants experiencing more competition and thus more interference when making a lexical decision on words in their L1. The authors contend that this demonstrates an inhibitory effect whereby the different-language prime word interferes in some way with the L1 target word recognition (Bijeljac-Babic et al., 1997, p. 453).

There have been similar findings with spoken-word recognition. Spivey and Marian (1999) conducted an eye-tracking experiment with Russian-English bilinguals in order to test whether they would experience parallel phonological activation of English words when they were listening and responding to instructions in Russian. Participants heard instructions in Russian telling them, for example to "pick up the stamp". The word for stamp is *marku* in Russian, and in addition to a stamp, participants also saw a marker on the table. Compared to the condition where the distractor object was not related phonologically to the target, the interlingual distractor condition (e.g. marku-marker) generated significantly more eye movements to the distractor. The phonological overlap between the Russian and English words, even though English was not used in the experiment, appears to have affected processing in the participants' L1.

Although there are quite a few differences, it is certainly possible to think of L2 processing in the same terms as L1, especially when we look at the recognition of utterances rather than single words. I am referring to context, and I imagine that L2 listeners can and often do use context to make up for imperfections in phonemic processing in the same way that an L1 listener uses it to understand homophones. For example, an L1 listener would not know if *roam* or *Rome* was the intended word when hearing /roʊm/, and in this case she would use context to resolve the ambiguity. In the same way, a learner of English who can not distinguish /r/ from /l/ could use the same skills in analyzing the context to determine if the word is *roam, Rome,* or *loam.* (Broersma & Cutler, 2008).

All of these studies detailing increased competition and thus slower word recognition in the L2 will come as no surprise to anyone who has struggled with learning a second language. Not only is the process of listening in the L2 hindered by activation of words from the L1, but the listener must also contend with the effects of inaccurate phonemic processing.

2.7 L2 Speakers with L2 Listeners

Non-native listening is complicated yet further when the spoken input comes from a non-native speaker. The complication arises due to potential production difficulties for L2 sounds by the speaker, compounded by similar perception difficulties for the listener as discussed above. There are, however, cases in which processing may be simplified. If both the speaker and the listener lack a phoneme in their respective L1, then variant pronunciation in the L2 may not cause the amount of processing difficulties that one would imagine. This has been shown to be the case in L2 listening to native speech (Cutler, Weber, Smits, & Cooper, 2004) as well as L2 listening to non-native speech (Weber et al., 2011). As one might expect, this has been demonstrated when both L2 listener and speaker share the same L1 (Bent & Bradlow, 2003). Their common L1 leads to the use of similar grammatical structures in the L2 as well as problems with similar sounds, making them more intelligible to each other than either one might be to a native speaker (Strange, 1995). A result that was more surprising in Bent and Bradlow's research (2003), although it is in line with my own experience in the language classroom, was that L2 listeners and speakers who did not share an L1 also demonstrated increased intelligibility when listening to each other in comparison with listening to a native speaker. The authors point out that the cause of this benefit might be due to shared grammatical and/or phonological structure of the participants' L1s, or to similarities in their respective interlanguages that are independent of their L1.

There is one large caveat to increased intelligibility among non-native speakers, however, and that requires that the variant pronunciation not create ambiguity by sounding like a different word. In other words, the string *he sinks at the ocean* may not activate /θɪŋk/ because /sɪŋk/ is an English word, unless of course the listener also produces /θ/ with the same variant pronunciation. Jenkins (2002) mentions numerous instances of miscommunication that she observed in her classroom research. For example, when one student made a presentation about /lɛt/ cars, the audience proceeded to ask questions about the daily rental fees when the presenter had intended to tell them about the color. In this case, experience with the accent in question may have facilitated comprehension: if the students in her classroom had heard a speaker with the same accent many times before, they may have had less difficulty comprehending the mix-up between /lɛt/ and /rɛd/, in much the same way that L1 listeners have been shown to adapt to a foreign accent (Bradlow & Bent, 2008; Norris et al., 2003).

In their study with German and Dutch L2 speakers of English, Hanulíková and Weber (2011) set out to examine the role of experience in spoken word recognition. They demonstrated that experience with accented pronunciation appears to facilitate recognition of variants in the L2. When they speak of "experience", they refer to their logical assumption that a Dutch speaker of English is likely more acquainted with Dutch-accented English, through hearing her own accent and that of her compatriots, than she is with other foreign accents. Using eyetracking, they showed that German listeners looked preferentially at English target words that began with /θ/ when they heard the same words pronounced with an /s/ in place of the /θ/, a substitution typical of German-accented English. The same was true for Dutch speakers when they heard a /t/ substituted for the /θ/. Interestingly, this facilitation occurred regardless of whether the /s/ or /t/ variant was produced by a German speaker or a Dutch speaker. The authors also undertook a perception experiment and determined that /f/ was more easily confused with /θ/ than /s/ or /t/ for both Dutch and German listeners. However, neither group of listeners experienced the same degree of looking preference for /θ/ words that were pronounced with an /f/ substitution. Also revealingly, they did not find a direct correlation between the way each *individual* participant pronounced /θ/ in a production study and the looking preferences for that participant. The participants in their study were living in their home countries, and therefore one can assume that they had more experience with their compatriots' pronunciation of English than with the accent of the other group. Hanulíková and Weber (2011) propose that L2 listeners who are resident in their home countries may have encountered the accented variant, in this case /sɪŋk/, before encountering the canonical form /θɪŋk/. However, the discrepancy between participants' production and processing of /θ/ as revealed in the eye-tracking experiment raises an important question about the relationship between production and processing, a question I seek to address in this study.

Weber, Broersma and Aoyagi (2011) found similar experiential effects in the processing of accented English by Dutch and Japanese speakers, but unlike Hanulíková and Weber (2011), they also found facilitation for foreign-accented words that were perceptually confusable with the standard pronunciation. This contrast is due to the specific sounds that they investigated. Hanulíková and Weber (2011) looked at /θ/ in relation to /f/, /s/, and /t/, all sounds which exist in both German and Dutch. In contrast, Weber et al. (2011) had Japanese participants listen to Dutch-accented English in which /æ/ was replaced by /ɛ/. As neither of these two sounds exists in Japanese, it is logical to assume that Japanese listeners might experience facilitation for the Dutch-accented variant, as they were likely unable to distinguish the difference between the variant form and the standard form.

Both Hanulíková and Weber (2011) and Weber et al. (2011) suggest that experience with accented English has an effect on recognizing words carrying the accent. This has been shown to be the case in various L1 studies, and my own research sets out to see whether similar experiential effects can be demonstrated with participants from different language groups, in this case Persian and Chinese, while also assessing the relationship between an individual's production and processing of variant forms. I will also examine Hanulíková and Weber's (2011) finding that /θ/ substitution variants produced priming effects even when the speaker was essentially mimicking a different accent than her own.

2.8 Face Priming

In addition to exploring the linguistic factors that determine processing of accented English, this study seeks to investigate a non-linguistic factor of potential importance by using a masked prime image of a face. Face priming has been used extensively in social psychology research to investigate ethnic and racial stereotypes, among other topics (e.g. Banse, 2001; Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Williams, & Powell, 2000). For example, Eberhardt, Goff, Purdie and Davies (2004) used masked primes of Black and White faces to investigate racial stereotypes in the United States. They found that in comparison with White face primes, Black face primes reduced the amount of time White participants needed to detect crime-relevant objects. They posit that the association in White Americans' minds between Blacks and crime served to create "perceptual processing biases" (Eberhardt et al., 2004, p. 879) that affected the detection of the target objects. If facial images of an ethnic group can activate a concept such as crime, it is possible that ethnically distinct facial primes may also affect how participants process accented spoken language.

2.9 Conclusion

In this chapter I have reviewed three areas of psycholinguistic research that form the backdrop for the present study: general research on spoken-language processing; research on variation and accents in the L1; and similar research in the L2. While a number of recent studies have begun to investigate L2 processing of L2 speech (Hanulíková & Weber, 2011; Weber et al., 2011), these have used participants from a relatively small number of language groups. In addition, to the best of my knowledge there are no studies that assess whether masked face primes have an effect on processing accented spoken language. This study is an attempt to increase our knowledge of non-native speech processing both from a linguistic and non-linguistic viewpoint.

Chapter 3 Methods

This chapter includes an explanation and discussion of the design and methods deemed most appropriate to address the research questions introduced in Chapter 1. This study involved two experiments, a speech production experiment and a cross-modal priming experiment, and each of them is described separately. In addition to a description of the methods of data collection, a justification of each method is provided, as well information about what was done to ensure the validity and reliability of the study. This is followed by an overview and justification of the linear mixed model approach used in the data analysis.

The written, electronic, and auditory data in this study were categorized and stored anonymously using a number for each participant. A summary of the study was submitted to NSD (*Norsk Samfunnsvitenskapelig Datatjeneste*) and acknowledgment was received (REF # 33300) that the study is not subject to the *Personopplysningsloven* or Personal Data Act.

3.1 Experiment 1: Production of /θ/

Common substitution variants for /θ/ for Chinese and Iranians have been documented (Chang, 2001; Wilson & Wilson, 2001). I confirmed these published accounts by speaking with two Chinese and two Iranian acquaintances before beginning this project. Nevertheless, in order to have a controlled basis for establishing the personal substitution preferences of both the Chinese and Iranian participants in this study, a production experiment was undertaken.

3.1.1 Materials and Design

The text read by the participants was a short story in English (see Appendix 1), and it was the same text as that used by Hanulíková and Weber (2011). The text included 56 instances of words with orthographic <th>. Nineteen of the words included a word-initial voiceless /θ/, while eight had a medial /θ/, and another six had /θ/ at the end of the word. The 24 instances of voiced /ð/ were repeated occurrences of the following words: *the, that, there,* and *than.* Five of the /θ/ words in the text also occur in the cross-modal priming study. I opted to elicit the production data before the cross-modal priming experiment to avoid any influence the accented variants in that experiment might have on the participants' production of the same and similar words. It is of course possible that there was an effect in the opposite direction, and that production of these items may have altered the participants' response times to the same items when they appeared as targets in the cross-modal priming study by "pre-priming" them.

3.1.2 Procedure

The participants were told that they would read a text into a microphone. They were also told not to worry if they did not know the meaning of all the words in the text. They were instructed to read at a normal speaking rate. When possible, I left the room during the recording. If this was not possible, I went to the other side of the room and faced away from the participants in order to limit their nervousness while recording the text.

After all the recorded data had been collected, I listened to the recordings and made a perceptual judgment regarding the pronunciation of the initial or final /θ/. Admittedly, this method of analysis had the potential to suffer from experimenter bias because I knew when listening to the recordings who was Iranian and who was Chinese. To ensure the reliability of my analysis, six of the 24 recordings were also analyzed by two native speakers of American English (three randomly selected recordings each) who had no knowledge of the topic of the experiment. They were asked to categorize the substitutions as falling into one of four categories: /θ/, /t/, /s/, or other. Analyst A listened to one Chinese and two Iranian recordings, and the percentage of overall agreement with my analysis was 80%, with a kappa inter-rater reliability of .71. Analyst B listened to two Chinese and one Iranian recording, and the percentage of overall agreement was 75%, with a kappa of .62.

The percentage and type of substitutions for /θ/ for each group of participants are listed in Table 4.1 in the next chapter. In addition, the number of substitutions for each participant was converted into a percentage score, indicating the degree to which that participant used variant pronunciation of /θ/. For example, a Chinese participant who produced 10 of the 25 /θ/ words in the English text with an /s/ substitution would receive a score of 40%, and this would be used in the statistical analysis to determine possible correlations between production and performance.

The validity of this type of production experiment is debatable. The main issue is something common to all experiments, which is that participants likely vary their pronunciation when they are in an experimental setting, especially when the researcher is a native-speaker and the overt topic of the experiment is the English language. Even though great care was taken to obscure the fact that /θ/ was the object of interest in this study, the simple fact that participants knew their English skills were being "tested" in some way likely led them to try their best to produce canonical pronunciations of all words in the text. Therefore, the production data recorded in this experiment may or may not be representative of the actual substitution frequencies for the participants. The only way around this problem would be to surreptitiously record two non-native speakers during a casual exchange in English, something nearly impossible and potentially unethical.

3.1.3 Participants

Twelve native speakers of Persian and twelve native speakers of Mandarin Chinese participated in this experiment. Three of the twelve Iranian participants and four of the Chinese participants were male. Participants were volunteers, and they received a gift card for food at the university in exchange for their participation. All of the participants reported having normal or corrected-to-normal vision, normal hearing, and no language related disabilities. The median age of the Iranian participants at the time of testing was 27 years, and the median for the Chinese participants was 26 years of age. One of the 24 participants had lived in the United Kingdom for 18 months, one in Singapore for two years, and three others had lived for less than six months in the United States; none of the others had ever lived in an English-speaking country. All of the participants had been living in Norway between one and three years at the time of the experiment, and they all stated that they use English in Norway on a daily basis. However, their formal English instruction took place in their home countries, and it is therefore likely that they heard accented-English for most of their lives.

After the cross-modal priming experiment, all participants completed the Lexical Test for Advanced Learners of English *LexTALE* (Lemhöfer & Broersma, 2012), a lexical-decision test used as an indicator of English proficiency in advanced learners. The Chinese participants had an average score of 64% correct, while the Iranian participants scored 59% on average. This is somewhat lower than the average scores of the participants in Weber et al. (2011), where the Dutch and Japanese participants scored 74% and 63% correct respectively.

After the LexTALE test, all participants filled out a written questionnaire (see Appendix 2) where they assessed their own English ability in reading, writing, speaking, and listening. There were four category options for each skill: beginner, intermediate, advanced, and fluent. Their self-assessment choices were converted into percentages (25%, 50%, 75%, 100%), and these were added to the data set.

3.2 Experiment 2: Cross-modal Priming

This experiment involved a subliminal masked visual prime image of a face, followed by an auditory prime, which was then followed by a visual target. Participants made a lexical decision on this final written string. Priming works on the following principle: when we hear or see a written word, our mental representation of that word is activated (Fernandez & Cairns, 2011), in addition to other words that share form and/or meaning with the prime word. The priming effect is the "residual activation from previously experienced stimuli" (Fernandez & Cairns, 2011, p. 191). Cross-modal priming was originally developed to study the effects of sentence context on lexical activation of ambiguous words (Swinney, 1979). It has since been used by countless researchers to investigate lexical access in real time as it exploits "one of the most robust lexical effects there is: people respond faster and more accurately when they process something they have just processed before" (Broersma & Cutler, 2008, p. 26). It is ideally suited to a study on variation in spoken language as it allows the participants to hear accented speech while making a lexical decision on canonical written forms, thus highlighting possible differences in the way variant forms are processed and stored in the mental lexicon.

The auditory prime and visual target portion of this experiment is modeled on the design of Weber et al.'s (2011) experiment with Japanese and Dutch learners of English. The listeners in the current experiment heard an auditory prime (e.g. /sɜrməl/) which was followed by the appearance of a visual target word on the screen (e.g. THERMAL). The participant then decided if the visual target was a real English word or not by pressing one of two buttons on a button box. When the auditory prime is related in form or meaning to the target, response times have been shown to be faster (Fernandez & Cairns, 2011).

Next I will describe the visual face prime, followed by the auditory prime and visual target word part of the experiment. The time sequence of the experiment can be seen in Figure 1.

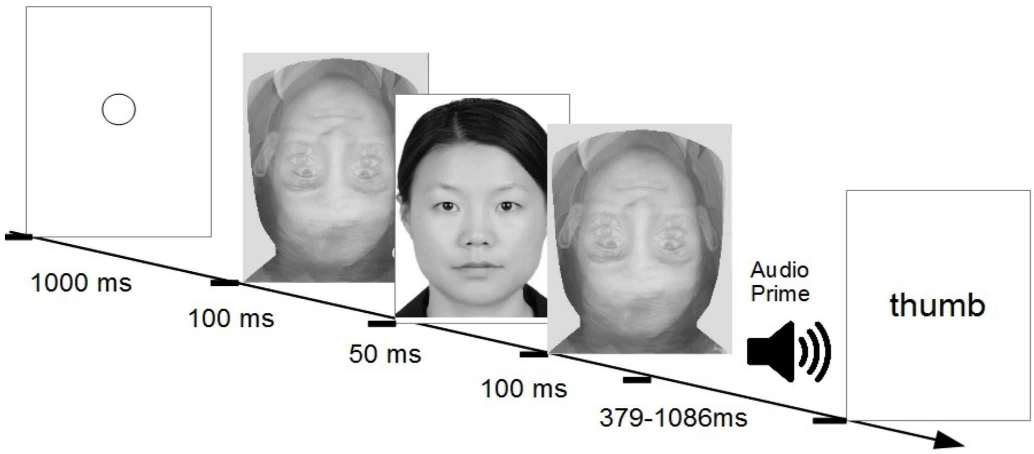


Figure 3.1. Time sequence of the priming procedure.

3.2.1 Materials and Design

3.2.1.1 Masked Face Prime

Facial images in priming experiments have been used extensively in social psychology to investigate gender and racial stereotypes, among other topics (e.g. Banse, 2001; Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Williams, & Powell, 2000). For example, Eberhardt, Goff, Purdie, and Davies (2004) subliminally primed participants with Black, White, or no faces. The White participants' task was then to indicate at what point they could identify a degraded image that slowly came into view. Some of the images were crime related while others were not. They found that in comparison with White face primes, Black face primes reduced the amount of time needed to detect crime-relevant objects. They posit that the association in White Americans' minds between Blacks and crime served to create "perceptual processing biases" (Eberhardt et al., 2004, p. 879) that affected the detection of the target objects.

To my knowledge, facial primes have not been used to assess if similar effects on linguistic processing could be demonstrated in a study of accented second-language speech. To this end, I developed this study to investigate whether facial images that match the accented speech heard in an audio prime would have a facilitatory effect on participants' response times to visual target words.

Four photographs of faces from various university student and faculty directory websites were used as subliminal prime images. Two photographs of Chinese adults (one male and one female) and two of Iranian adults (one male and one female) were selected for their similarity in having neutral expressions, a lack of eyeglasses and facial hair, and facial characteristics that fit a rough stereotype for each nationality. The backgrounds on the photographs were standardized using Microsoft Paint, and they were all reduced to a size of 360 by 500 pixels using Google Picasa photo editor. The images were approximately 70mm wide by 100mm high when viewed on the 17" screen of the Dell laptop computer used to conduct the experiment with all participants.

A mask image was created by morphing all four photographs into one, with two of them inverted. After the 1000ms duration fixation point (o) in the center of the screen, a mask image was displayed, followed by the prime image and another mask. The two masks were displayed for 100ms each, and the prime face was displayed for 50ms. The 50ms prime duration was decided on by comparison with similar research (Banse, 2001; Eberhardt et al., 2004) and with a pre-test with two volunteers (18- and 26-year-old Norwegian females) who did not take part in the main experiment. A very short face prime of 10.5ms was used by Banse (2001), and when this image was masked, they found reverse priming effects. Eberhardt et al. (2004), on the other hand, found clear priming effects for masked face images displayed for 30ms. I decided that in order to ensure priming effects, it would be best to display the face primes for the longest time possible that was still below the threshold of conscious recognition. The monitor refresh rate is a limiting factor in determining the possible durations of prime images. The refresh rate on the computer used in this experiment is 60Hz, meaning that images can only be displayed for multiples of 16.66ms (i.e. 16.6ms, 33.3ms, 50ms, 66.6ms, etc.). I ran ten trials from the experiment with two volunteers who did not know that the flash at the beginning of each trial included facial images. I began by using a display length of 34ms for the facial image while keeping the two masks constant at 100ms each. I then showed them a printed paper with 12 faces, four of which appeared in the experiment (see Appendix 3). The eight distractor faces included four images that fit the same nationality and gender as the experimental face primes (i.e. one male and female image of both Chinese and Iranians). The remaining four faces were neither Chinese nor Iranian, although two were male and two were female. Neither at 34ms nor at 50ms were the volunteers able to recognize any of the four experimental faces from the list of twelve printed faces. However, at a 67ms duration, one of the volunteers accurately pointed to three of the experimental images. I thereby determined that 50ms was just below the threshold for supraliminal recognition of facial images in one of these volunteers, and that this display duration would likely be appropriate for the experiment. That a 50ms display rate was below the perceptual threshold was confirmed during the main experiment: only one participant out of 24 accurately recognized two of the experimental images, and two other participants recognized one image from the post-test (Appendix 3). Although this demonstrates that 50ms was fast enough to avoid conscious perception by the participants in this study, it may be the case that it was too fast for some or all of the participants to subliminally register the images they were seeing. This possibility will be discussed in Chapter 4.

3.2.1.2 Auditory Prime and Lexical Decision

Twenty-four English words were chosen as visual targets, 12 containing /θ/ in word-intial position, and 12 containing /θ/ in word-final position (see Appendix 4). The words had an average lemma frequency of 62 per million in the Corpus of Contemporary American English (Davies, 2008). Of the words with initial /θ/, six were monosyllabic and six were disyllabic; eleven of the /θ/-final words were monosyllabic, and one was disyllabic. From an original list of over 40 /θ/ words, these 24 were selected because they had the highest frequency among words where both the /s/ and the /t/ substitution resulted in a pseudoword. In other words, a word such as *thank* was not used, even though it has a high frequency, because both /sæŋk/ and /tæŋk/ are real words. If substitution for /θ/ results in an alternate word, it is then not possible to determine if a participant is responding to an accented variation of *thank* or to the real words *tank* or *sank*, in this case.

Each target word was preceded by one of three possible auditory prime words: a Persian-accented prime, a Chinese-accented prime, or an unrelated prime. The accented primes were variations on the standard pronunciation of the word such that the /θ/ was replaced with a /t/ in the Persian version and with an /s/ in the Chinese version. Mandarin Chinese speakers have been shown to most often replace /θ/ with /s/ (Rau, Chang, & Tarone, 2009), while Persian speakers more often opt for /t/ as a substitute (Wilson & Wilson, 2001). This substitution tendency was also confirmed by the production data from Experiment 1 in this study. It is important to note that while neither Chinese nor Persian has the /θ/ phoneme, both languages do have /s/ and /t/. In addition, /θ/ words produced with a canonical pronunciation were not included as primes, as the combination of canonical and non-canonical pronunciations of the same words within one experiment has been shown to affect processing of the non-canonical items (Kraljic, Samuel, & Brennan, 2008b).

The unrelated, or control primes were matched for frequency and were both semantically and phonologically unrelated to the target words. The unrelated primes had an average lemma frequency of 60 per million in COCA (Davies, 2008). While only 24 control primes were needed for the experiment, a total of 36 potential words were recorded by a female native speaker of Persian and a male native speaker of Mandarin Chinese. I then selected the prime words that had the fewest obvious accent markers and substitutions by both the Persian and Chinese speakers.

In addition to the 24 /θ/ target words, the experiment included a set of non-/θ/ words that were matched with the /θ/ words for frequency and syllable length (see Appendix 5). These functioned as both a control experiment and as fillers, and they had an average lemma frequency of 69 per million in COCA (Davies, 2008). They were preceded by one of three possible real-word auditory primes in the same way as the experimental items: a Chinese-spoken prime identical to the target, an Iranian-spoken identical prime, or an unrelated control prime balanced between speakers. Items in this control experiment were chosen in a similar way to the control primes mentioned above: 36 potential words were recorded by both speakers, and I then selected those with the least accented pronunciation. Therefore, in contrast to the experimental /θ/ words, the accented primes that preceded these targets were not purposely accented on a specific phoneme.

The additional filler trials included 24 English words and 72 pseudowords as targets. The 24 English word filler targets had an average lemma frequency of 77 per million in COCA (Davies, 2008). They were preceded by various types of pseudoword primes: 8 minimal pairs that differed at the beginning of the word; 8 minimal pairs that differed at the end; and 8 pseudowords that were phonologically unrelated. The 72 pseudoword targets were preceded by various prime types: 24 of them had real English word primes (8 word-initial minimal pairs, 8 word-final minimal pairs, 8 unrelated); 24 of them had identical pseudoword primes; and 24 had pseudoword primes that fit the model above of 8 word-initial minimal pairs, 8 word-final minimal pairs, and 8 unrelated pseudowords.

All of the auditory primes were recorded twice, once by a female native speaker of Persian and once by a male native speaker of Mandarin Chinese. The recording sessions for each speaker took place on different days. The experimental items were first recorded without instructing the speakers to intentionally produce an /s/ or a /t/ in place of the /θ/ in the words, in hopes that they would do so naturally. However, as the experimental items were in a list, it quickly became obvious to the speakers that the topic of interest was the /θ/ sound, and they made a noticeable effort to produce what they thought would be a "correct" /θ/. The Persian speaker is a linguistics student who is more proficient in English than the Chinese speaker, and therefore her approximation of /θ/ was more accurate, in my opinion. I then gave them a list of the items with modified spelling such that the word *thief* was written as "tief" and "sief". I did not inform the speakers that they should try to approximate a Persian or a Chinese accent, but rather simply asked that they substitute /t/ and then /s/ for /θ/ in each of the words. Both speakers produced the primes in citation style. The items were recorded in a soundproof room using a Shure KSM44 microphone and were then saved on a computer at a sample rate of 41.5 kHz. Each audio prime was then extracted from the recording using the sound-editing software Praat (Boersma & Weenink, 2012).

Six versions of the experiment were constructed so that each of the three prime conditions was spoken by each of the two speakers, thus leading to a factorial design of three (prime condition: Chinese /s/, Persian /t/, control) by two (prime accent: Chinese or Persian). Each participant began with six practice trials that included both real word and pseudoword trials, followed by the 24 experimental trials and 120 filler trials in random order. Twelve of the experimental items contained /θ/ at the beginning of the word, and twelve had /θ/ in the word-final position. Each experimental target word appeared once in each version of the experiment, with eight experimental items in each of the three prime conditions (Persian /t/, Chinese /s/, unrelated control), and these were further subdivided such that four were spoken by the Persian speaker and four by the Chinese speaker. Each participant saw 24 real-word targets and 24 pseudoword targets with identical auditory primes, 24 real-word targets and 16 pseudoword targets with unrelated auditory primes, and 24 real word targets and 32 pseudoword targets with primes that were similar but not identical to the target. In the case of the real-word targets, the similar primes were the accented versions mentioned above, while the pseudoword primes were minimal pairs, half of which differed at the beginning of the word and half at the end. In total, participants saw the same number of real word and pseudoword targets: 72 of each.

3.2.2 Procedure

Participants completed the experiment in a quiet room, either alone or with the researcher sitting as far away as possible. After completing the production experiment described above, they read the instructions on the computer screen (see Appendix 6). The instructions stated that they would see a circle followed by a flash, and they would then hear a voice and see a sequence of letters. If the sequence of letters was a real English word, they were instructed to press the green button. If the sequence of letters was not a real English word, they were to press the red button. The participants then put on the closed headphones and completed the six practice trials with the researcher in the room. A brief synopsis of the instructions then appeared again on the screen. After confirming that they did not have any further questions, the researcher left the room when possible.

The experiment was designed and controlled using Paradigm 2.0 software (*Paradigm 2.0*, 2012). The button box used was a Cedrus RB-530, and the headphones were Sony MDR-XB300. Response times were recorded from the onset of the visual target word stimulus.

The stimulus onset asynchrony varied depending on the length of the auditory prime. The auditory primes ranged from 379ms to 1086ms in length, with a mean of 672ms and a standard deviation of 113ms. The mean length of all the auditory primes recorded by the Chinese speaker was 616 ms, while the corresponding mean for the Iranian speaker was 727ms. This difference between the speakers was evident in the 24 experimental primes as well, where the means for the Chinese and Iranian speakers were 651ms and 761ms respectively.

3.2.3 Participants

The participants were the same as those in Experiment 1: 12 native speakers of Persian and 12 native speakers of Mandarin Chinese.

3.3 Methods of Data Analysis

All data were analyzed using SPSS and R (R Core Team, 2012) and the R packages language R (Baayen, 2009) and lme4 (Bates, Maechler, & Bolker, 2012). The data were analyzed using linear mixed-effects models. In order to avoid the "language-as-fixed-effect fallacy" detailed by Clark (1973), participants and target words were used as crossed random effects (Baayen, 2008). The fixed effects were the language of the participant, the spoken language of the prime, and the prime condition (/s/, /t/, control). The additional covariates of age, gender, handedness, English self-assessment, percentage of substitutions for /θ/ from the production study, and LexTALE score were also used in constructing the best-fitting model.

Mixed-effects models have recently been shown to offer advantages in the analysis of repeated measures data in psycholinguistic research (Baayen et al., 2008). Since Clark's (1973) seminal paper, it has been common in psycholinguistic research to compute separate F ratios of between-group variance over within-group variance for both participants and items. This has been seen as a way to account for the random variation among participants and the fact that the linguistic materials used in any one experiment are a sample of the "unbounded combination of finite lexical items" (Baayen et al., 2008, p. 390). Advances in statistics over the last 30 years, however, have made it possible to combine random effects and thereby generalize across both participants and items within a single model (Winter, 2011). According to Baayen et al. (2008), although mixed-effects models are not widely known in linguistic research, they are now used in many other branches of science, medicine, and engineering.

There are additional advantages to be gained by analyzing data using mixed-effects models. First, it is possible to test the combined effect of two independent variables, one of which is numerical while the other is categorical (Winter, 2011). Baayen et al. (2008) also point out that while counterbalancing is often used to offset fatigue or learning effects over the course of an experiment, mixed-effects models allow these longitudinal effects to be brought into the statistical model. Specifically, using by-subject random slopes for the trial order takes into account variation among participants due to fatigue and learning effects.

In addition, mixed-effects models make it possible to predict responses on categorical dependent variables, such as the yes/no responses in this study. Although ANOVA is commonly used for categorical data analysis in psycholinguistics, Jaeger (2008) argues that even when proportional data are transformed using arcsine-square-root, this type of analysis can lead to spurious results. He argues that ANOVA is particularly problematic when a participant in an experiment performs at or near ceiling, with the proportion of correct answers at or near 1 (Jaeger, 2008, p. 441). Jaeger goes on to assert that mixed-effects models have a number of advantages over ANOVA in the analysis of categorical data: they limit overfitting of a model by using penalized likelihood in place of absolute likelihood; they have greater power to detect true effects; and they allow testing to determine if hypothesized random effects such as participants and items should be included in the model, rather than simply using them because it has become standard procedure in the field (2008, p. 444).

After finding a model that provides the best fit to the data, there are some checks that must be performed. Using visual inspection of a plot of residuals (see Appendix 7) against fitted values, the model should be checked for normality and homogeneity (Winter, 2011). If there is no linear trend observable in the plot of residuals, this indicates that the error of the model is not systematic and hence acceptable in terms of homogeneity and normality. In addition to analysis of the residuals, each potential model with both random and fixed effects should be compared to a null model that contains only random effects. Models that do not differ significantly from the null model are then rejected (Winter, 2011).

The p-values listed throughout this thesis were validated with Markov chain Monte Carlo simulations (Baayen et al., 2008). For models that contain by-subject or by-item random slopes for other predictors, t-values from the random-slope model will be presented together with MCMC p-values from the corresponding random-intercept model.

The mixed-effects model package in R allows one to include a lengthy number of covariates when constructing models and assessing them for goodness of fit to the data. This provides powerful analytical potential, but it also introduces a threat to the internal validity of the study by increasing the likelihood that unseen confounds or highly correlated variables will lead to spurious conclusions. For example, this study included four self-assessment measurements of English ability from participants in addition to their LexTALE scores (Lemhöfer & Broersma, 2012). Therefore, there was a need to assess the independence of these variables from each other in the analysis phase. To that end, the correlation tables from each model were examined during the model-building process, and any variables that were correlated above .30 were either decorrelated by means of creating a ratio of the two, or removed from the analysis if it was determined that they were confounded with another variable. Specific examples will be discussed in the next chapter.

Chapter 4 Results & Discussion

This chapter provides a presentation and analysis of the results from both experiments, with links to each of the four research questions and references to relevant previous research. The first section deals with the production experiment, followed by discussion of the cross-modal priming experiment. Possible reasons for some of the unexpected results are given, as well as suggestions for future research.

4.1 Results of Experiment 1: Production of /θ/

The categorization of the participants' pronunciation of the 25 /θ/-items from the English text are detailed in Table 4.1 below. Although most speakers used one substitution consistently, some speakers produced two or sometimes three different substitutions, with the pattern heavily dependent on the location of the /θ/ (word-initial or word-final) and the surrounding vowels and consonants. For example, some speakers who were able to closely approximate a native-like production of /θ/ in the word-initial position in a word such as *thief*, had more difficulty when producing the /θ/ in *depth*, often resulting in an /s/ substitution. Both Chinese and Persian speakers are reported to have difficulty producing clusters of two, and especially three consonants at the beginning or end of a word because these are non-existent in those languages (Chang, 2001; Wilson & Wilson, 2001).

Table 4.1 Percentages of word-initial and word-final /θ/ substitutions by participant group when reading text in Appendix 1 (percentages are rounded up and numbers of occurrences are in parentheses)

Pronunciation of /θ/ in word-initial position (19 words)				
	/θ/	/s/	/t/	Other
Participants				
Chinese	77% (176)	15% (34)	6% (14)	2% (4)
Iranian	54% (123)	0%	46% (105)	0%
Pronunciation of /θ/ in word-final position (6 words)				
	/θ/	/s/	/t/	other
Chinese	50% (36)	44% (32)	0%	6% (4)
Iranian	76% (55)	14% (10)	6% (4)	4% (3)

The first research question addressed by this study was whether participants' production of English /θ/ would correlate with their performance on a lexical decision task that included accented production of the same phoneme. I will first discuss the frequencies of substitution variants for /θ/ in each speaker group together with a look at possible correlations between individual pronunciation and performance on the lexical decision task.

4.1.1 /θ/ Pronunciation among Chinese Participants

It is interesting to note the difference in the percentage of /s/ substitutions for word-initial (15%) and word-final (44%) /θ/ among the Chinese participants. Granted that there were only six /θ/-final words in the experiment, but the difference is still noteworthy. The large difference is due to five participants who produced almost all of the six /θ/-final words with an /s/ substitution. These same participants were responsible for some of the /s/ substitutions for /θ/-initial words, but those substitutions were more evenly distributed.

As shown in Table 4.1, the Chinese participants produced 77% of the words with initial /θ/ as /θ/, and 15% with an /s/ substitution. A full 82% of the substitutions are attributable to two speakers, and it is therefore not informative to discuss which words received the most /s/ substitutions. Those two participants, 17 and 29, substituted 79% and 68% respectively of all the /θ/-initial words in the text with /s/. One might be tempted to attribute such pronunciation difficulty to lower-than-average English proficiency. However, in this case it is interesting to note that participants 17 and 29 both scored slightly above the mean for Chinese participants (64%) on the LexTALE test, with 69% and 66% correct respectively. Their scores on the /θ/ items in Experiment 2 were 96% and 100% respectively, compared with the mean for Chinese participants (96%) and the mean for all participants (95%). Therefore, in this case there is a slight disjuncture between production and proficiency as gauged by two lexical decision tests.

In addition to the expected /s/ substitutions, there was one participant who was responsible for 13 of the 14 /t/ substitutions for word-initial /θ/. After finishing the experiment, this participant asked, as did many others, what the purpose of the experiment was. After explaining that I was investigating the processing and production of /θ/, the participant said that she had learned while studying in Singapore to substitute /t/ for /θ/; she understood this substitution to be the most "proper", and she had consciously used a /t/ substitution when reading the text for the production experiment. It was unclear to me whether she was able to produce the canonical /θ/ or not.

Regarding the word-final /θ/, the Chinese participants produced 50% of them with /θ/ and 44% with /s/. Ninety-one percent of the /s/ substitutions are from five of the twelve Chinese participants. The same two participants who had difficulty with word-initial /θ/, 17 and 29, are among the five. The other three participants, on the other hand, had none (participants 4 and 30) or very few substitutions (3 substitutions for participant 28) in word-initial position. Consonant clusters are noted to be difficult for Chinese speakers of English, and this is especially true at the ends of words (Chang, 2001).

One way to address the research question regarding possible correlation between production and performance of /θ/ is to examine the RTs of individuals with a high percentage of substitutions. In the case of the Chinese participants, there were two who displayed high levels of /s/ substitution for /θ/ as described above: participants 17 and 29. At first glance, it would appear that the RT data does not support the hypothesis that people who produce more /s/ substitutions show facilitation when listening to similar substitutions: both participants were slower when responding to /s/ primes than /t/ primes, and in the case of participant 17, the RTs for the /s/ prime condition were even 72ms slower than the control condition. However, there are a number of problems with this analysis. First, a sample of two participants is much too small, and as this was a repeated measures design, the items in question for each prime condition are different, making this type of comparison impossible. Second, the average RTs for all Chinese participants also show this tendency: slower responses overall in the /s/ prime condition. The third and most important problem with this analysis becomes clear when the RTs for Chinese participants are separated by the speaker of the prime word as seen in Table 4.2. There was a 167ms difference in RTs across all Chinese participants between /s/ primes spoken by the Chinese speaker and those spoken by the Iranian speaker. This was the only prime condition in which such as large difference is evident, and it points to the possibility that both the type of substitution, /s/ or /t/, and the accent of the speaker were important for Chinese listeners. Whether or not this difference is statistically significant will be discussed below in the model analysis. Therefore, the analysis of RTs for the two Chinese participants is inconclusive regarding a possible correlation between production and performance in this study. Below I will analyze the effect of the speaker on processing of the prime words, look at the level of accentedness for /θ/ words produced by all participants, and determine if this data set provides any evidence for the hypothesized correlation.

4.1.2 /θ/ Pronunciation among Iranian Participants

In contrast to the Chinese participants, the Iranian participants had more difficulty with /θ/-initial words, producing 46% with a /t/ substitution. They also differed in that /t/ substitutions were not confined to just a few speakers but were more widespread. Indeed, there was only one Iranian participant who did not produce any /t/ substitutions. The word that posed the greatest difficulty for the Iranian speakers was *threshold,* with 10 out of 12 participants substituting /t/ for /θ/. Persian does not have consonant clusters within single syllables, and Persian speakers are known to have difficulty producing /θr/ (Wilson & Wilson, 2001). I imagined that this was the cause of their difficulty; but while the two tokens of *throughout* saw a /t/ substitution in 7.5 out of 12 participants, the second and third words with the highest /t/ substitution did not have a consonant following the /θ/ but rather a vowel: *thousand* with nine /t/ substitutions and *thoughts* with eight. I then surmised that the cause of their difficulty might lie with the surrounding vowel sounds. If these were unfamiliar, they may have led to difficulty with the initial consonant. However, according to Wilson and Wilson (2001) the /aʊ/ in *thousand* and the /ɔː/ in *thoughts* both have near equivalents in Persian and do not often cause difficulty for speakers. I anticipate that the low frequency of *threshold,* and therefore the likelihood that participants were unfamiliar with the word, led to a high rate of /t/ substitution. This is supported by the fact that a more familiar word such as *three*, which occurred four times in the text, saw an average of only five /t/ substitutions. There may also be other factors affecting pronunciation such as word length, the final sound of the previous word, and/or difficulties with other sounds that occur later in the same word, such as /ʃ / in *threshold*.

The word-final items, in contrast, showed more canonical tokens (76%) of /θ/ and fewer substitutions among Iranian participants. Interestingly, of the 17 substitutions for word-final /θ/, 10 of them were with /s/. The literature on common substitution patterns for Iranian speakers of English states that /t/ is a common substitute for /θ/, but it does not specify the word position (Wilson & Wilson, 2001). Based on this admittedly small sample, it would appear that /s/ is preferred to /t/ in word-final position. Although their research only looked at /θ/ in word-initial position, Hanulíková and Weber (2010) also found differential substitution patterns among German and Dutch speakers of English.

4.1.3 Discussion of Production Experiment

The elicitation of English pronunciation data by having participants read a short text was useful for two reasons. First, it demonstrated that the documented /θ/ substitution preferences for Chinese and Persian speakers apply to the participants in this study: although the percentages vary, the preferred substitution for /θ/ among the Chinese participants is /s/, while the Iranians more often use /t/. Second, the data showed that overall, the frequency of /θ/ substitution among the participants in this study was quite low: 22% for Chinese and 36% for Iranian participants. The fact that the vast majority of /θ/ words were produced with something near a canonical pronunciation must be taken into account when analyzing any potential correlation between production and perception.

While the production data are useful, there are some problems with them as well. One issue is something common to all experiments, which is that participants likely vary their pronunciation when they are in an experimental setting, especially when the experimenter is a native-speaker and the overt topic of the experiment is the English language. Even though great care was taken to obscure the object of interest in this experiment, the simple fact that participants knew their English skills were being "tested" in some way likely led them to try their best to produce canonical pronunciations of all the words in the text. The Chinese participant mentioned above who had lived in Singapore is a case in point. Therefore, the production data recorded in this experiment may or may not be representative of the actual substitution frequencies for the participants.

A second problem is due to the use of different words in the production and cross-modal priming experiments. This was a conscious choice in order to avoid possible frequency effects. However, as described above for the word *threshold*, there are obvious differences in substitution frequencies on the level of individual words: some /θ/ words may elicit frequent substitutions due to neighboring sounds or familiarity, whereas others may not. Therefore, it is possible that the substitution frequencies observed in the words in the production experiment are not representative of, and do not correspond with, the frequency of stored variants for the words in the cross-modal priming experiment.

A third issue involves the validity of this experiment. A sample size of 12 participants from each language group is sufficient to insure internal validity for the repeated measures design of the cross-modal priming experiment. However, it is less than adequate in order to generalize about the substitution preferences of the larger population of Chinese and Persian speakers. Ideally, one would need to sample non-native speakers who are resident in their home countries and who have never lived abroad. Although even that sample would likely display variation in substitution preferences depending on the participants' exposure to native pronunciation of English, the accent of their non-native English teacher, and individual differences in their ability to acquire an L2.

Be that as it may, after looking at the distribution of substitutions averaged across participants from the same language group, as well as additional research that documents the English /θ/ substitution preferences for Persian and Chinese speakers (Chang, 2001; Wilson & Wilson, 2001), it is reasonable to assume that the participants in this study differ in their past experience with hearing substitutions for English /θ/. In other words, the Chinese have heard more /s/ substitutions and the Iranians have heard more /t/ substitutions for /θ/. If this past experience influences how easily participants can recognize such substitutions, then Iranian participants are expected to experience a facilitatory effect when hearing /t/ substitutions, while Chinese participants should experience a similar effect when hearing /s/ substitutions. This was investigated in the cross-modal priming experiment documented below.

4.2 Results of Experiment 2: Cross-modal Priming

The cross-modal priming experiment was designed to address research questions number one and two in this study: (1) Does production of /θ/ correlate with performance on a lexical decision task that includes accented production of the same phoneme?; and (2) Does experience with the accented English from one's home country facilitate word recognition of similarly accented English? I hypothesized that both groups of participants would respond faster and with fewer errors when hearing the substitution that matched their accent (/s/ for Chinese and /t/ for Iranians), and that individuals who produced more substitutions in the production study would display this trend even more than those with fewer accented tokens of /θ/.

In this section I will review the results from this experiment and discuss some potential reasons for the unexpected results. The results from the experimental /θ/ items will be discussed first; this includes a mixed-effects model for the entire data set, separate models for the data from the Chinese and Iranian participants, and a binomial model to examine the error rates. This will be followed by a similar set of models to explain the results from the control experiment, which included non-/θ/ items that were matched for frequency with the experimental /θ/ items.

The trimmed data set from the experiment was analyzed using linear mixed modeling, and the best-fitting models and their output results will be presented in each section below. All data were analyzed using SPSS and R (R Core Team, 2012) and the R packages language R (Baayen, 2009) and lme4 (Bates et al., 2012). In order to avoid the "language-as-fixed-effect fallacy" detailed by Clark (1973), participants and target words were used as crossed random effects (Baayen et al., 2008). The main fixed effects were the language of the participant, the accent of the prime, and the prime condition (/s/, /t/, control). Additional covariate fixed effect factors were added to the models as they were run and criticized for goodness of fit.

Using visual inspection of a plot of residuals (see Appendix 7) against fitted values, I checked for normality and homogeneity of the models (Winter, 2011). The p-values listed throughout this thesis were validated with Markov chain Monte Carlo sampling with 10,000 replications (Baayen et al., 2008).

Table 4.2 shows the mean RTs for correct responses with a logarithm of RT below 7.5, on the 24 /θ/ experimental items measured from the onset of the visual target, and the error percentage for Chinese and Iranian participants listening to the recordings of the Chinese speaker and the Persian speaker. The RT data suggest clear priming effects for the Iranian participants when listening to both speakers and for the Chinese participants when listening to the Chinese speaker. However, the slower RTs for Chinese participants to the /s/ primes spoken by the Persian speaker in comparison to the control primes were not expected. Some possible explanations for this apparent inhibitory effect will be discussed below.

Table 4.2. Chinese and Iranian participants' percentage of incorrect responses and mean RTs of correct responses for /θ/ target words in Chinese (/s/), Iranian (/t/), and control prime conditions, separately for primes spoken by Chinese and Persian speaker.

			% Error		RTs in ms (std)	
			Participant		Participant	
			Chinese	Iranian	Chinese	Iranian
Chinese Speaker	Prime Condition	/s/	2.2	8.9	687 (192)	855 (296)
		/t/	6.4	0	786 (331)	826 (246)
		control	6.5	2.5	789 (196)	938 (322)
Persian Speaker	Prime Condition	/s/	4.3	8.9	854 (346)	805 (321)
		/t/	0	9.1	732 (245)	798 (273)
		control	4.3	5.1	738 (235)	898 (283)

4.2.1 Model Design and Criticism: /θ/ Items

All incorrect responses were removed, equivalent to 6.8% of the experimental trials. The response times (RTs) were log transformed to approximate a normal distribution and reduce skewness. Trials with RTs above 7.5 (1808ms or 1.88 SD above the mean) were excluded from the analysis, resulting in the removal of 5.4% of the correct-response data. This equates to the removal of 11.8% of the experimental trials. Mixed-effects models were then fitted to this trimmed data set.

In addition to using the natural logarithm of the RTs, other variables were transformed as well. After an initial test model, it was found that the self-assessment values for English speaking ability and listening ability were highly correlated. This is not surprising, as a high self-assessment on one would likely signal a high score on the other. Therefore, it was decided to combine these two scores by creating a new variable *EnglishRatio*, which is a participant's English speaking ability divided by her listening ability. For example, a participant with a speaking score of 100% and a listening score of 50% would now have an *EnglishRatio* score of 2.

The prime and the target items for each trial were matched as closely as possible for frequency, resulting in a strong correlation between these two variables in the analysis. Therefore, these were combined in a way similar to the English scores above by dividing the natural logarithm of the target word frequency by the logarithm of the prime frequency to create a new variable. As with the RTs, using the logarithm helps to normalize the data, and in the case of frequency values it weights the differences between less-frequent words more heavily. For example, among the 24 experimental /θ/ items in this experiment, the frequencies ranged from 7 per million for *thermal*, to 341 per million for *health* (Davies, 2008). If we were to compare these words with others that occur 14 and 348 times per million, these would both be 7 occurrences higher, respectively, than *thermal* and *health.* However, the word that occurs 14 times per million is double the frequency of *thermal*, whereas the same can not be said for the word with a frequency of 348 and *health*. Using the logarithm of these values helps to take this fact into consideration in the data analysis.

Other variables were also transformed. The variable *Trial* is a record of the order in which each trial was presented to the participants. The trials were randomized, so each participant was presented with the trials in a different order. The order of presentation has been shown to present effects of familiarization in some participants and fatigue in others (Baayen & Milin, 2010; Baayen, 2008). Baayen (2008) points out the potential for spurious correlation when using a trial variable that is bounded by zero to the left. He explains that the solution is to center the variable by subtracting the mean from each value. In this study, that means that trials 1 to 144 are now labeled as trials -72 to 72, with 0 as the mean. This transformed *Trial* variable is labeled *cTrial* below. Figure 4.1 shows the change in RTs for each participant during the 144 trials in the experiment, with a clear effect of familiarization for participants 15 and 26, and possible fatigue in the case of participants 24 and 32.

In order to test whether by-subject random slopes for trial presentation order are justified, I first fit the following mixed-effect model in R:[2]

```
>lmer1 = lmer (log(RT) ~ cTrial + (1|SUBJECT) + (1|ITEM), data_correct)
```

This is a random-intercept model where *RT* is the dependent variable, *cTrial* is a fixed effect and SUBJECT and ITEM are crossed-random effects. SUBJECT refers to the participants in this study, and ITEM refers to the target words. This model calculates a mean effect of trial order for all participants, taking into account different intercepts, or mean starting points, for each participant and item.

2 Note that the following two models are presented purely for explanatory purposes; the rationale behind this experiment is that I think the language of the participant, the accent of the prime, and the prime condition are important predictors of RT, not simply the trial presentation order. Because these two models only look to see the effect of trial order on RT, they overestimate its effect to some degree.

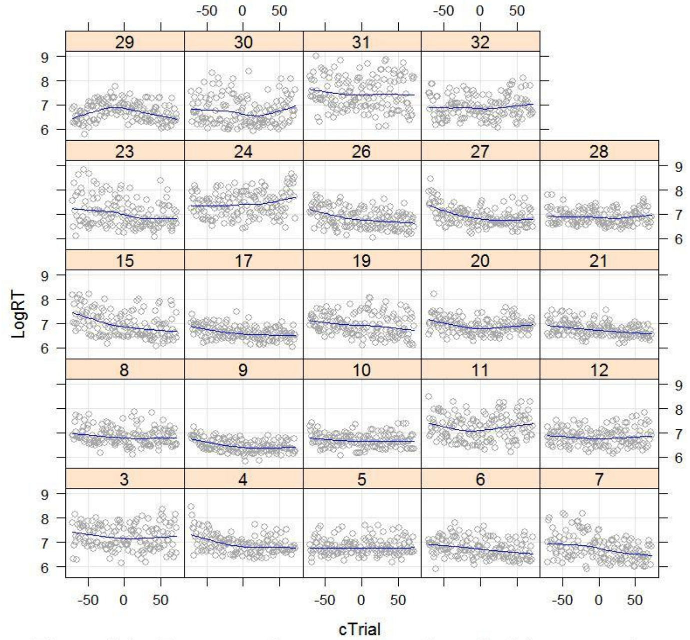


Figure 4.1. Response times as a function of trial presentation order for each participant.

The output from this model shows a significant main effect for *cTrial* (t=-5.31). If visual inspection of Figure 4.1 had shown that all participants had increased in speed over the course of the experiment, model *lmer1* would likely be justified because it would compensate for this familiarization by subtracting a specified number of milliseconds from each trial. However, because the plots show that most participants got faster but a few others got slower, model *lmer1* likely does not provide the best fit to the data. Not only would it mistakenly subtract from the RTs of participants who got slower, but it would underestimate the degree to which the majority got faster. Therefore, I constructed a second model with by-subject random slopes for *cTrial*, thus taking into account that participants showed differing effects with respect to trial presentation order.

```
>lmer2 = lmer (log(RT) ~ cTrial + (1+cTrial|SUBJECT) + (1|ITEM), data_correct)
```

The output from *lmer2* showed that the main effect for *cTrial*, while still significant, is lower (t=-3.70). The next step involved testing *lmer1* against *lmer2* by conducting an analysis of variance (ANOVA). The results were significant (p<.01), and the residual value was lower for *lmer2*, indicating that it explains more of the variation in the data set. Therefore, the inclusion of by-subject random slopes for trial presentation order is justified, and they will remain in the model. I have gone into detail here in order to exemplify the process that was used throughout the analysis phase with mixed-effects models: one predictor variable or covariate is added or removed, and the new model is tested against the previous one using ANOVA to determine if the change was significant and if the inclusion or removal provides a better fit to the data.

The best-fitting model for the data set from both the Chinese and Iranian participants is shown in Table 4.3. Participants and items were included as crossed-random effects, and participant language, accent of the prime, and prime condition (/s/,/t/,or control) were the main fixed-effect predictors. By-subject random slopes for *cTrial* were included as described above.

Table 4.3. Results from the statistical analyses of /θ/ items for Chinese and Iranian participants combined

	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	7.8357	0.3306	23.70	0.0001
Participant Iranian	0.1304	0.0702	1.86	0.1202
Accent Persian	-0.0718	0.0492	-1.46	0.1996
Prime /s/ vs. control	-0.1657	0.0493	-3.36	0.0044
Prime /t/ vs. control	-0.0508	0.0491	-1.03	0.3532
cTrial	-0.0016	0.0004	-4.08	0.0001
Participant LexTALE Score	-0.0132	0.0039	-3.34	0.0044
Self Assessment Ratio - English Speaking/Listening	-0.3730	0.1707	-2.19	0.0116
Participant Iranian:Accent Persian	0.0451	0.0730	0.62	0.6166
Participant Iranian:Prime /s/ vs. control	0.0720	0.0713	1.01	0.4788
Participant Iranian:Prime /t/ vs. control	-0.0450	0.0712	-0.63	0.3144
Accent Persian:Prime /s/ vs. control	0.2644	0.0690	3.83	0.0010
Accent Persian:Prime /t/ vs. control	0.0024	0.0686	0.03	0.9362
Participant Iranian:Accent Persian:Prime /s/ vs. control	-0.2858	0.1011	-2.83	0.0090
Participant Iranian:Accent Persian:Prime /t/ vs. control	-0.0409	0.1006	-0.41	0.8204

Note: Estimates indicate the regression coefficients for the fixed effects; p-values were obtained by Markov-Chain Monte Carlo (MCMC) sampling with 10,000 replications.

There was a significant main effect (*t*=-3.34; *p*MCMC =0.004) for the score on the LexTALE test (Lemhöfer & Broersma, 2012), such that participants with higher scores responded faster. The participants' self assessment of their English ability expressed as a ratio of listening over speaking skill also showed a significant (*t*=-2.19; *p*MCMC=0.012) main effect.

Using the results from the production experiment, a variable was created to represent the strength of each participant's accent on /θ/. It is the percentage of the /θ/ words produced with the expected accent for that participant's language group (i.e. /s/ for Chinese and /t/ for Iranians). This variable was not found to be a significant (*t*=0.03) predictor of RT. It was also eliminated due to the skewed results from the production experiment described above in section 4.1: a large percentage of the substitutions for /θ/ were produced by a small number of participants with strong accents. In addition, the distinction between items with initial /θ/ and final /θ/ was found to be highly correlated with the frequency of the item, making that variable also unusable in the analysis. /θ/-final words had a mean frequency per million three times higher than /θ/-initial words.

The participants' age was not a significant predictor (*p*=0.309), nor was their gender

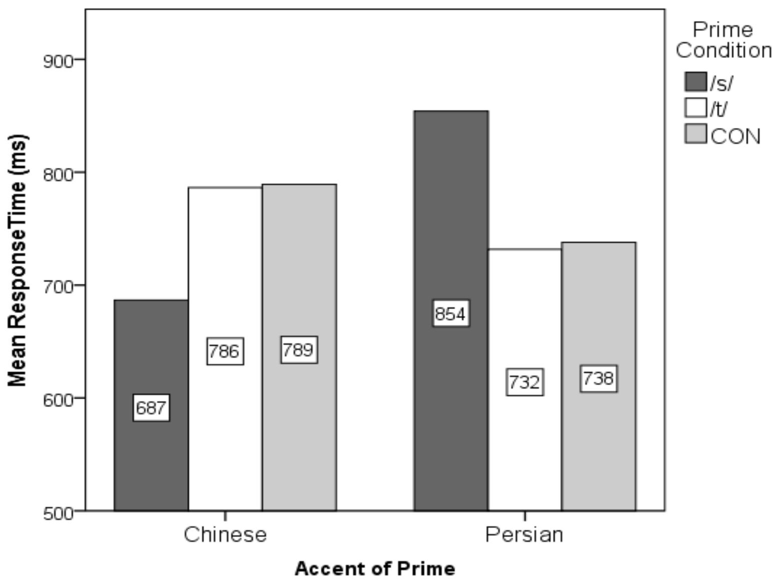
(p=0.766) nor handedness (p=0.633). Whether or not the subliminal face matched the accent of the spoken prime was not a significant (t=0.12) predictor of RT in this model or any subsequent model below.

In terms of the main predictors of participant language, prime accent, and prime condition, the results from this model are difficult to interpret due to the significant (t=-2.82; p=0.009) three-way interaction between those factors. Based on the mean RT data in Table 4.2 above, it appears that Chinese participants reacted quite differently depending on the speaker of the /s/ prime. Therefore, in order to determine if this difference is significant, it is necessary to divide the data and fit models to separate Chinese and Iranian data sets.

4.2.1.1 Chinese Participants: /θ/ Items

Figure 4.2 shows the mean RTs, measured from the visual target onset, for correct responses with a logarithm of RT below 7.5, in the three prime conditions for Chinese participants listening to the recordings of the Chinese speaker and of the Persian speaker.

Figure 4.2. Average RTs for correct responses for /θ/ items from Chinese participants after Chinese-accented (/s/), Persian-accented (/t/) and unrelated primes spoken by Chinese and Persian speakers.



The results from the Chinese participants were unexpected because not only do they differ from the Iranian participants, but they also appear to contradict a previous study where facilitation for substitution variants of /θ/ occurred regardless of whether the accent was authentic or mimicked (Hanulíková & Weber, 2011).

The best-fitting model for the data from the Chinese participants is shown in Table 4.4.

There was a main effect for trial presentation order (*t*= -2.15 ;*p*MCMC= 0.029), but by-subject random slopes for trial were not justified (*p*=0.37). This indicates that the Chinese participants as a group got faster throughout the experiment. There was also a marginally significant main effect for participants' self-assessment of their English ability (*t*= -1.95; *p*MCMC= 0.061). The degree to which participants substituted /s/ for /θ/ in the production experiment was not a significant (*t*=-1.35) predictor of RT.

Table 4.4. Results from the statistical analyses of Chinese participants' responses to /θ/ items

	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	7.3332	0.4778	15.35	0.0001
Accent Persian	-0.0749	0.0503	-1.49	0.1612
Prime /s/ vs. control	-0.1578	0.0500	-3.15	0.0028
Prime /t/ vs. control	-0.0523	0.0503	-1.04	0.3444
cTrial	-0.0008	0.0004	-2.15	0.0294
Self Assessment Ratio - English Speaking/Listening	-0.5011	0.2564	-1.95	0.0608
Accent Persian:Prime /s/ vs. control	0.2514	0.0704	3.57	0.0002
Accent Persian:Prime /t/ vs. control	0.0107	0.0705	0.15	0.8764

However, the estimates for the predictors that are most important for this analysis, accent of the prime and the prime condition, indicate a potential problem with this model. The fact that the main effect for the /s/ prime has a significant negative *t*-value (*t*=-3.15), while the interaction of the /s/ prime with the Persian accent shows a significant positive value (*t*=3.57) suggests that the model may be over-fitting the data. Therefore, further subdivision of the data into primes spoken by the Chinese speaker and primes spoken by the Persian speaker is justified. Application of this same model to the trials where the prime was spoken by the Chinese speaker showed significant priming effects (estimate = -0.1476; *t* = -3.18; *p*MCMC = 0.004) for the /s/ condition, and no significant priming effects for the /t/ condition (estimate=-0.0381; *t*=-0.81; *p*MCMC=0.502). This result fits with the hypothesis of this study that Chinese participants should experience facilitation when hearing the /s/ substitution because it is the common variant for /θ/ among Chinese speakers.

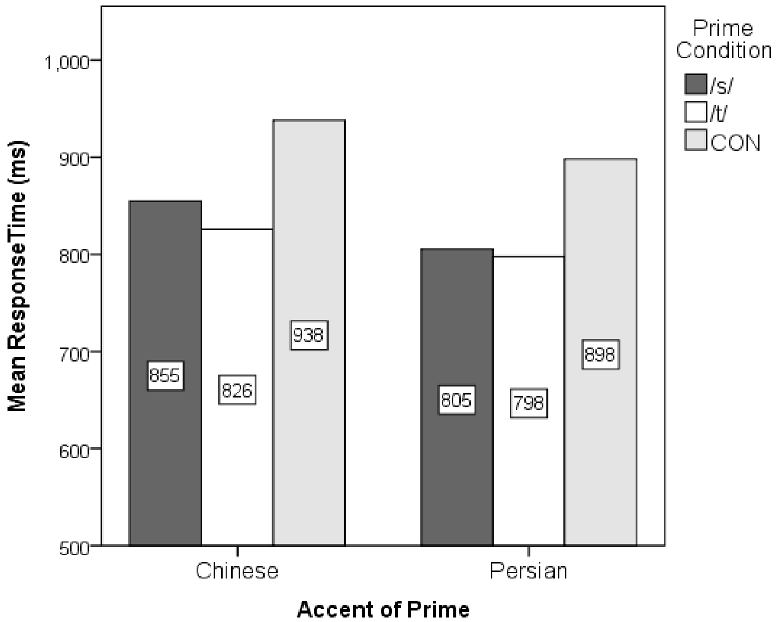
In contrast, when the model was applied to the trials where the primes were spoken by the Persian speaker, the results were somewhat unexpected. Although the /t/ condition showed a similar lack of priming effects (estimate=-0.0276; *t*=-0.52; *p*MCMC=0.692) in comparison to the control, the /s/ prime spoken by the Iranian speaker led to a marginally significant inhibition for Chinese listeners (estimate=0.1000; *t*=1.88; *p*MCMC= 0.0694). In other words, Chinese listeners only experienced facilitation for the /s/ substitution when it was produced by the Chinese speaker.

When the substitution did not fit with the context of the accent, they actually responded more slowly than they did to the unrelated control prime. That by-subject random slopes for prime condition were not significant (p=0.60) confirms that this pattern of facilitation and inhibition existed for all the Chinese participants to some degree or another. Before presenting possible reasons for these unexpected results, I will discuss the data from the Iranian participants and the control experiment.

4.2.1.2 Iranian Participants: /θ/ Items

Figure 4.3 shows the mean RTs, measured from the visual target onset, for correct responses with a logarithm of RT below 7.5, in the three prime conditions for Iranian participants listening to the recordings of the Chinese speaker and of the Persian speaker.

Figure 4.3. Average RTs for correct responses for /θ/ items from Iranian participants after Chinese-accented (/s/), Persian-accented (/t/), and unrelated primes spoken by Chinese and Persian speakers.



The best-fitting model for the data from the Iranian participants is shown in Table 4.5.

Table 4.5. Results from the statistical analyses of Iranian participants' responses to /θ/ items

	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	6.5520	0.4616	14.19	0.0001
Accent Persian	-0.0486	0.0287	-1.69	0.0998
Prime /s/ vs. control	-0.0991	0.0351	-2.82	0.0044
Prime /t/ vs. control	-0.1063	0.0358	-2.97	0.0006
cTrial	-0.0024	0.0006	-4.05	0.0001
Participant LexTALE Score	-0.0128	0.0037	-3.49	0.0166
PrimTargFreqRatio	0.9846	0.3953	2.49	0.0178

In contrast to the Chinese participants, there were no significant (t<0.79) interaction effects between the prime condition and the accent. The degree to which participants substituted /t/ for /θ/ in the production experiment was not a significant (t=0.36) predictor of RT. In addition, by-subject random slopes for trial presentation order were justified for the Iranian participants (p=0.005), indicating that they differed among each other more than the Chinese in terms of whether their RTs increased or decreased throughout the experiment.

There are two important differences between the Iranian and Chinese participants. First, the Iranian participants showed a moderately significant (t=-1.69; pMCMC=0.099) facilitatory effect when hearing their own accent, regardless of the prime condition. Second, the results show similar priming effects for /s/ substitutions (t=-2.82 ; pMCMC=0.004) and /t/ substitutions (t=-2.97; pMCMC=0.001), regardless of the speaker. Indeed, when the /s/ condition is taken as the reference level, the /t/ condition shows only a non-significant (estimate=-0.0187; *t=-0.38*; pMCMC=0.47) advantage over the /s/.

These results indicate a substantial difference in the way that Iranian and Chinese participants in this study responded to substitution variants for English /θ/. The prime condition played an important role for the Chinese participants: they experienced significant facilitation when the /s/ prime was spoken by the Chinese speaker, inhibition for /s/ primes spoken by the Persian speaker, and no priming effects for the /t/ substitution. The Iranian participants, on the other hand, appear not to have been affected by the substitutions for /θ/ but rather by the accent itself, at least to a marginally significant degree. In order to confirm whether these effects are actually due to the prime condition and/or accent, the error rates for the /θ/ items and the RTs for the control experiment will be analyzed next.

4.2.1.3 Error Rates: /θ/ Items

The error rates for Chinese and Iranian participants for items in each of the three prime conditions and with primes spoken by the Chinese and Persian speakers can be seen above in Figure 4.2. A linear mixed-effects model (binomial family using *lmer* in R) was fitted and tested on the error data, with participants and target words as crossed random effects and with the participants' language, the accent of the spoken prime, and the prime condition (/s/, /t/, control) as the fixed-effect predictors (e.g. Baayen et al., 2008). Whether or not the participant answered correctly was the dependent variable.

Of the 24 participants, 13 of them made either zero or one error on the experimental items. A binomial model where most participants performed at or near ceiling has very limited predictive power. Indeed, neither the accent of the prime, nor the prime condition, nor the self-assessment of English ability had a significant (*p*s>0.12) effect on error rates. The only significant predictors were the logarithms of the target frequency (p=0.012) and prime frequency (p=0.019). In other words, participants made more errors on less frequent words, which is to be expected.

4.2.2 Model Design and Criticism: Non-/θ/ Items

Non-/θ/ words were included in the cross-modal priming experiment as a control. The 24 items in this portion of the experiment were matched for frequency, and as closely as possible for syllable length with the 24 /θ/ words. Unlike the /θ/ items, the primes in this portion of the experiment were not altered on one specific phoneme; the only variation was due to the fine phonetic detail of each speaker's accent. Therefore, comparison with the data from the /θ/ items will help to show the relative contribution of substitution variants and overall foreign accent to the pattern of responses from each participant group.

The same criteria were used to trim the data for the non-/θ/, or control, experiment as were used for the /θ/ items. Similarly, all the same transformations were used on the variables. The mean RTs and error percentages are shown in Table 4.6 below. Table 4.7 shows the best-fitting model for the data from both the Chinese and Iranian participants for the non-/θ/ items.

By-subject random slopes for trial order were not justified for this data set (p=0.089). There were no significant three-way interactions between the participant language, the accent of the prime, and the prime condition (*t*s>-1.10). The two-way interactions between accent and prime type (t=-1.03) and participant language and accent (t=-1.48) also did not reach significance.

Table 4.6. Chinese and Iranian participants' percentage of incorrect responses and mean RTs of correct responses for non-/θ/ target words in identical and control prime conditions, separately for primes spoken by Chinese and Persian speaker

			% Error		RTs in ms (std)	
			Participant		Participant	
			Chinese	Iranian	Chinese	Iranian
Chinese Speaker	Prime Condition	identical	6.4	7.9	688 (234)	790 (274)
		control	4.3	6.4	784 (287)	928 (304)
Persian Speaker	Prime Condition	identical	5.3	6.6	714 (257)	724 (256)
		control	2.1	11.4	789 (251)	953 (307)

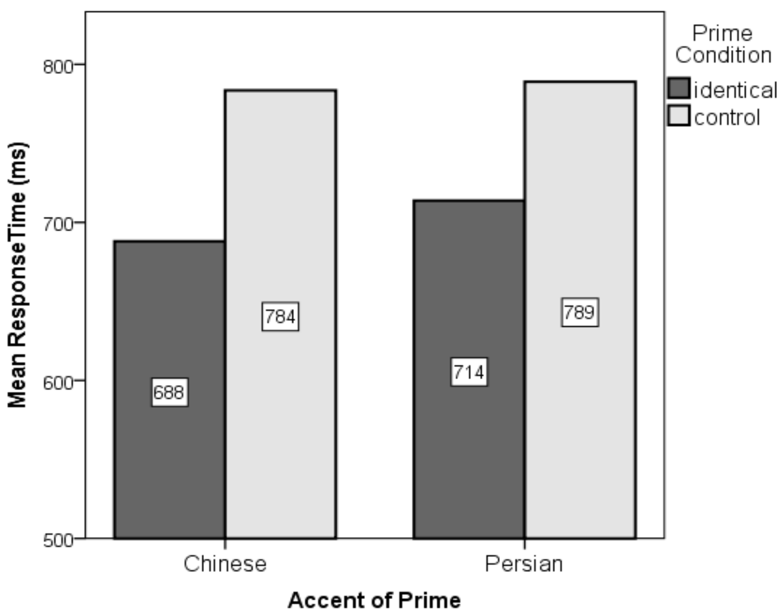
Table 4.7. Results from the statistical analyses of Chinese and Iranian participants' responses to non-/θ/ items				
	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	7.9080	0.3029	26.11	0.0001
Prime identical vs. control	-0.1686	0.0240	-7.03	0.0001
cTrial	-0.0014	0.0003	-4.90	0.0001
Self-assessment ratio - English speaking/listening	-0.4320	0.1692	-2.55	0.0102
Participant LexTALE score	-0.0129	0.0037	-3.52	0.0004

In addition, there was no main effect for participant language (*t*=1.50); in other words, when there was no substitution variant as there was with the /θ/ items, both groups behaved quite similarly, showing significant priming effects for the identical prime regardless of the accent (*t*=-7.03; *p*MCMC=0.0001). However, the *t*-value (*t*=-1.47) for the interaction between participant and accent hints at the results from the main experiment where Iranians showed marginally significant facilitation when hearing their own accent. For that reason, and for the sake of consistency, I will subdivide the data set and fit models to the Chinese and Iranian data separately.

4.2.2.1 Chinese Participants: Non-/θ/ Items

Figure 4.4 shows the mean RTs measured from the visual target onset, for correct responses with a logarithm of RT below 7.5, in the two prime conditions for Chinese participants listening to the recordings of the Chinese speaker and of the Persian speaker.

Figure 4.4. Average RTs for correct responses for non-/θ/ items from Chinese participants after identical and unrelated control primes spoken by Chinese and Persian speakers.



A preliminary model for the data from the Chinese participants responding to the non-/θ/ items is shown in Table 4.8 below, and the final model in Table 4.9.

Table 4.8. Preliminary results from the statistical analyses of Chinese participants' responses to non-/θ/ items

	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	6.9632	1.0414	6.69	0.0001
Accent Chinese	-0.0141	0.0301	-0.47	0.6434
Prime identical vs. control	-0.1271	0.0378	-3.36	0.0026
cTrial	-0.0008	0.0004	-2.02	0.0434

Table 4.9. Final results from the statistical analyses of Chinese participants' responses to non-/θ/ items

	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	6.6239	0.0537	123.34	0.0001
Prime identical vs. control	-0.1208	0.0317	-3.81	0.0002
cTrial	-0.0008	0.0004	-2.09	0.0000

There was no significant interaction between the accent and the prime type (t=-0.10), although in this case the prime type was a binary distinction between identical and control. In stark contrast to the results from the experimental items, Chinese participants showed significant (t=3.36; *p*MCMC=0.003) facilitatory priming effects regardless of the accent of the prime. Unlike the Iranian participants, they showed no significant facilitation when hearing their own accent (t=0.47; *p*MCMC=0.643). These results help to clarify the findings from the experimental items that showed an inhibitory effect when Chinese participants heard the /s/ prime condition spoken by the Persian speaker. Clearly, it was not the Persian accent itself that disrupted processing, but the fact that the /s/ substitution occurred in an unexpected context.

4.2.2.2 Iranian Participants: Non-/θ/ Items

Figure 4.5 shows the mean RTs, measured from the visual target onset, for correct responses with a logarithm of RT below 7.5, in the two prime conditions for Iranian participants listening to the recordings of the Chinese speaker and of the Persian speaker.

The results from the control experiment also help to clarify the pattern that emerged with the /θ/ items in which the Iranians showed a marginally significant facilitatory effect when hearing their own accent, regardless of whether the substitution was an /s/ or a /t/. Indeed, the results from the non-/θ/ control items, shown in Table 4.10 below, confirm this pattern.

Figure 4.5. Average RTs for correct responses for non-/θ/ items from Iranian participants after identical and unrelated control primes spoken by Chinese and Persian speakers.

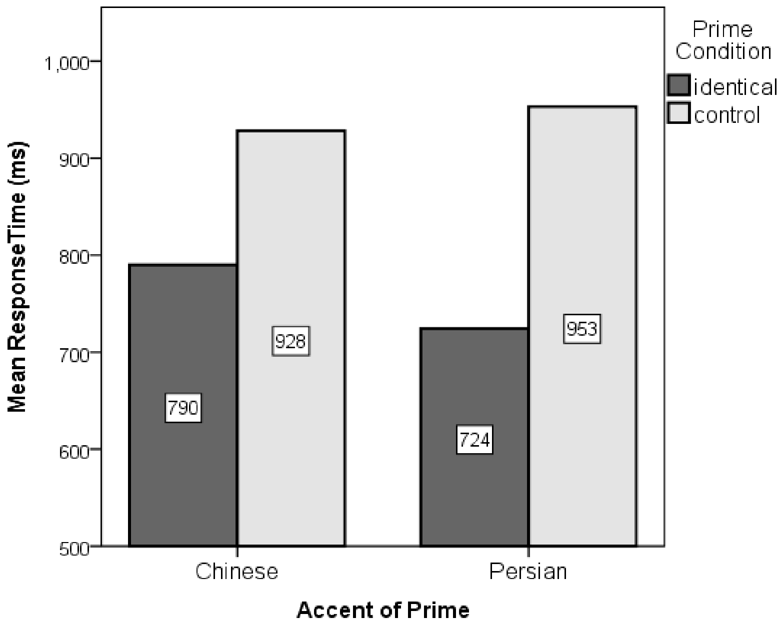


Table 4.10. Results from the statistical analyses of Iranian participants' responses to non-/θ/ items				
	Estimate	Std. Error	*t* Value	*p*MCMC
(Intercept)	7.4403	0.2293	32.43	0.0001
Accent Persian	-0.0587	0.0330	-1.78	0.0898
Prime identical vs. control	-0.2191	0.0348	-6,28	0.0001
cTrial	-0.0020	0.0005	-3,74	0.0001
Participant LexTALE Score	-0.0101	0.0038	-2,65	0.0304

By-subject random slopes for trial order were justified in this model (*p*MCMC=0.036), and there was no significant interaction effect between accent and prime type (*t*=-1.20). While the bar graph of the mean RTs for Iranian participants in Figure 4.5 shows a noticeable difference in priming effects, this turns out to be only a marginally significant (*p*MCMC =0.089) facilitation when hearing their own accent.

4.2.2.3 Error Rates: Non-/θ/ Items

The error rates for Chinese and Iranian participants for items in the identical and control conditions for primes spoken by the Chinese and Persian speakers can be seen above in Table 4.6. A linear mixed-effects model (binomial family using *lmer* in R) was fitted and tested on the error data, with participants and target words as crossed mixed effects and with the participants' language, the accent of the spoken prime, and the prime condition (identical or control) as the fixed-effect predictors (e.g. Baayen et al., 2008). Whether or not the participant answered correctly was the dependent variable.

Of the 24 participants, seven of them made either zero or one error, while eight of them made two errors on the non-/θ/ items. Neither the accent of the prime, nor the prime condition, nor the language of the participant had a significant (ps>0.23) effect on error rates. The only significant predictors were the ratio of the participants' self-assessed English listening ability over speaking ability (p=0.036), the logarithm of the target frequency (p=0.057), and the logarithm of the prime frequency (p=0.041). In other words, participants who ranked their English listening skills as higher than their speaking skills made fewer errors, and all participants made more errors on less frequent words, which is to be expected. The errors were so heavily concentrated that 51% of them occurred on just three target words: *boulder, choke,* and *rip*.

4.3 General Discussion

The results from the production and cross-modal priming experiments point to a complex interplay between accents and substitution variants among the non-native English speakers in this study. I will now address each of the research questions that was presented in Chapter 1 based on the results above.

4.3.1 Question 1: Correlation Between Production and Performance

In response to research question one of this study, it was not possible to determine whether a correlation existed between the participants' production of /θ/ and their performance on the cross-modal priming experiment. As was mentioned in section 4.1 above, the production experiment results for the Chinese participants were highly skewed, with a mean of 22% (standard deviation 29) substitution of /s/ in the 25 /θ/-words in the text. The results for the Iranian participants were closer to being normally distributed, with a mean of 36% (standard deviation 22) substitution of /t/ for /θ/. However, the variable created to represent this information for each participant was not found to be a statistically significant predictor of RTs or error rates for either group of participants. There are a number of potential reasons for this outcome.

First, participants showed varying levels of substitution for /θ/. While there were five Chinese participants who produced all 25 words with a canonical /θ/, the remaining seven Chinese and all twelve of the Iranians produced some, but not all, of the words with /s/, /t/, or other substitutions. If some participants had produced 100% of the items with a substitution, and the remaining participants had produced 100% of the items with canonical /θ/, it would be possible to conclude that /θ/ substitution is simply a black and white matter of articulatory ability. However, the results point to a more complex picture where the phonology of the speaker's L1, the position of the /θ/ in the word, its neighboring sounds, and the frequency of the word, among other possible factors, all come into play to determine how it is pronounced. Therefore, the use of a single variable to represent a participant's level of accent was simply not a sufficiently fine-grained measure. Future research on this topic would require testing participants on the same /θ/ words in both the production and the priming experiments in order to rule out some of the confounds mentioned above.

A second possible reason for the lack of a significant result may simply be that the sample size was too small. A larger sample would likely have provided more variation and potentially more possibilities to observe a correlation.

4.3.2 Question 2: Experiential Effects

Question two asked whether participants would respond more quickly and with fewer errors when they heard /θ/ words with substitutions typical of the accent from their home country. Such experiential effects are seen to be linked to the statistical frequency of the variant forms: the more often one has heard an /s/ substitute for /θ/ for example, the stronger the link will be between the variant and canonical representations of the word (Ranbom & Connine, 2007). Both Hanulíková and Weber (2011) and Weber et al. (2011) found experiential effects for substitution variants with German and Dutch, and Japanese and Dutch participants respectively. I hypothesized that I would find similar effects for the Chinese and Iranian participants in this study.

Indeed, the prime condition played an important role for the Chinese participants. They experienced significant facilitation when the /s/ prime was spoken by the Chinese speaker. These results confirm the hypothesis of this study, and suggest that linguistic experience plays a role in non-native listening. Significant facilitation for the more-frequent variant form, /s/ in this case, suggests that they are linked more strongly to the underlying lexical representations than the less-frequent /t/ variant (as in Ranbom & Connine, 2007).

However, the remaining results were not so straightforward. While they showed facilitation for /s/ primes spoken by the Chinese speaker, the Chinese participants displayed an inhibitory effect for /s/ primes spoken by the Iranian speaker, and no priming effects for the /t/ substitution. If the differing effects were due to general characteristics of the Iranian accent, then I would have expected to see a similar facilitatory and inhibitory pattern for Chinese participants in the control experiment. This was not the case. It appears that the mimicked nature of the /s/ prime from the Iranian speaker caused the inhibition, and this will be discussed below in response to question number three.

The Iranian participants, on the other hand, appear not to have been affected by the substitutions for /θ/ but rather by the accent itself, at least to a marginally significant degree. This pattern was confirmed by the results from the control experiment, where both Chinese and Iranian participants showed priming effects for words without segmental substitutions, and Iranians continued to show marginally significant facilitation when hearing their own accent.

It is tempting to contrast these results with the findings from Hanulíková and Weber (2011), in which participants showed similar priming effects for variant forms regardless of the speaker of the prime. However, in their study, the accent of the prime was a between-subjects factor such that each group of participants only heard the variant forms (/s/, /t/, or /f/ for /θ/) produced by one speaker. If I had followed a similar design, it is possible that the participants in this study would have displayed a similar tendency. The Iranian participants in particular may have shown facilitation for the /t/ prime over the /s/ prime had they heard only one speaker or the other. However, when hearing two accents, one of which was their own, it appears that facilitatory priming effects were maximized simply by the accent.

The results for the Iranian participants are not particularly surprising. Even though the Iranian speaker's pronunciation may have differed markedly from native pronunciation, the Iranian listeners were well prepared to interpret her speech. They share the same L1 and thus likely have similar accents when they speak English, resulting from transfer of sounds and stress patterns. In an off-line study where participants with various L1s listened to English sentences read by speakers with the same and different L1s, Bent and Bradlow (2003) found this to be the case, and they called it the "matched interlanguage speech intelligibility benefit" (p. 114).

Although facilitation when hearing one's own accent is not surprising, the complete lack of priming effects for the /t/ substitution for Iranian participants is difficult to explain. The Iranian participants in this study produced 36% of the /θ/-words with a /t/ substitution in the production experiment, and 3% with an /s/. This degree of /θ/ substitution is actually higher than that found by Hanulíková and Weber (2011) for Dutch participants, who substituted /t/ for /θ/ only 29% of the time when reading the same English text. Therefore, it is reasonable to assume that the Iranian participants favor the /t/ substitution for /θ/, and that priming effects should occur. However, this was not the case. It appears that when hearing primes from both speakers, the familiarity of the Persian accent was strong enough to supersede any increased facilitation for /t/ primes over /s/ primes. As mentioned above, one possible way to explore this question in future research would be use the accent of the prime as a between-subjects factor, and use a larger sample size.

An alternative explanation for these results could be that the inclusion of items with word-final /θ/ weakened the priming effects. Half of the 24 experimental items included word-initial /θ/

and half were word-final. According to some models of spoken-language processing, word-initial segments may be more important than segments that occur later in the word (Connine, Blasko, & Titone, 1993). If this is the case, Iranian participants may have recognized the word-final /θ/ items before reaching the /s/ or /t/ substitution at the end of the word, thus rendering the substitution meaningless.

4.3.3 Question 3: Authentic vs. Imitated Accent

The two speakers in this study were not asked to imitate the accent of the other speaker group. When the prime stimuli were recorded, neither speaker knew that the /s/ and /t/ variations would be used to approximate the accent of another non-native speaker group. However, the Persian speaker certainly did know that /t/ is a common substitution for Persian speakers of English because she is the one who informed me of that fact. Nevertheless, the substitution variants that did not match the reported preference for each speaker's L1 can be said to be "imitated" in this study. Therefore, the inhibitory priming effect seen for Chinese participants when listening to the Iranian speaker's /s/ primes can be seen as a reaction to an imitated accent.

The research on imitated vs. authentic accents is limited. Neuhauser and Simpson (2007) investigated the ability of native German listeners to judge the authenticity of accents produced by native and non-native German speakers. They found that listeners were not particularly successful at judging when an accent was authentic or mimicked, but they concede that there was a large amount of variation depending on the speaker. Markham (2007) also found variation in native Swedish listeners' ability to accurately judge the authenticity of real and imitated dialectal accents in their L1. Both of these studies deal with accent authenticity of sentence-length utterances in the L1, and are therefore not directly applicable to this study. However, it is reasonable to imagine that if native speakers are unable to judge the authenticity of foreign and dialectal accents in their L1, non-native speakers would be even less likely to be successful at this task. Granted, the Chinese participants in my study were not asked to judge the authenticity of the accented prime words. But the inhibitory effect seen when they heard /s/ primes spoken by the Persian speaker signals that, in some way, they heard these items as not authentic.

Hanulíková and Weber (2011) found subtle differences in the /s/ and /t/ produced by German and Dutch speakers of English when they were asked to produce similar accented versions of /θ/ words, but these did not affect comprehension for listeners. They hypothesize that the participants in their study may have had sufficient experience with L2 English spoken by people with different L1s to permit short-term perceptual learning (as in Bradlow & Bent, 2008).

In the case of the Chinese and Iranian participants in this study, it is possible that their accents were simply more different from each other than the German and Dutch accents in the above study. But if this were the case, then the Chinese would likely have displayed significantly slower RTs in the control experiment when listening to the non-/θ/ primes produced by the Persian speaker. Therefore, there must be something special about the combination of the /s/ prime and the Persian accent as heard by the Chinese participants. It is not possible to know whether the effect comes from the /s/ phoneme itself or its substitution for /θ/. This is due to the fact that the control experiment primes were specifically chosen not only because they matched the frequency of the /θ/ items, but also because they did not begin with /θ/, or in most cases with /s/ or /t/. But there was one prime used in the control experiment that began with /s/: *seek* was the control prime for the target *church.* If the inhibitory effect for Chinese participants when listening to /s/ substitutions produced by the Iranian speaker is due to the /s/ itself and not its substitution for /θ/, then I would expect to see a similar pattern in trials where the prime *seek* was spoken by the Iranian speaker. However, this does not appear to be the case. In trials with *seek* as the prime, Chinese participants responded with mean RTs of 759ms to the Chinese-spoken prime, and 732ms to the Iranian-spoken prime. These RTs are well within one standard deviation of the mean RTs for Chinese participants to the Chinese-spoken primes (784ms) and the Iranian-spoken primes (789ms). Therefore, it appears unlikely that the inhibitory effect is caused by /s/ in its canonical form, but rather only when it is used as a substitute for /θ/. There are two obvious weaknesses of this analysis: *seek* is a control prime, meaning that participants responded not to *seek* but to the target *church*; and RTs for *seek* come from only two participants in each speaker condition, an admittedly small sample.

One possibility for future research would be to perform acoustical analyses on the /s/ and /t/ produced by both speakers when substituting for /θ/, and compare these with /s/ and /t/ that occur in canonical forms of words. By priming participants with both groups of words, one would then have a clearer picture of whether the inhibitory effect comes from the /s/ itself, or its substitution for /θ/.

With the results available from this study, I can only conclude that some fine phonetic detail in the /s/ sound when it was substituted for /θ/ by the Iranian speaker caused a disruption in lexical activation for the Chinese listeners. As was discussed in Chapter 3, the details of both vowel and consonant sounds are determined by the sounds that surround them, such that the /p/ in *wrap* is not the same as the initial /p/ in *paper* (McQueen, 2005). Some recent studies have sought to uncover the importance of such subcategorical differences by separately splicing the onset of a competitor word and a pseudoword onto the target word. For example, Dahan, Magnuson, Tanenhaus, and Hogan (2001) used eyetracking to assess the role of lexical competition in such subcategorical mismatches. They found that in both cases, eye movements were delayed in comparison to the canonical prime, but that splicing from the competitor word caused an even greater delay than splicing from a pseudoword. They argue that previous studies (Marslen-Wilson & Warren, 1994) where a difference between competitor- and pseudoword-spliced mismatches was not found suffer from a lack of detail resulting from their use of lexical decision tasks rather than eye tracking.

It may be possible to argue that the /s/ substitution for /θ/ in this study is roughly equivalent to the /s/ of a pseudoword as in Dahan et al. (2001). For example, the Iranian and Chinese speakers in this study were asked to produce the word *sermal,* (*thermal* with an /s/ substitution), which is arguably a pseudoword. Unlike in the above study, there were no competitors or distractors in this study, and the methods and goals were obviously different. However, it would appear that for the Chinese listeners, the /s/ in *sermal* was processed as a canonical exemplar of *thermal* when it was spoken by the Chinese speaker, and as a pseudoword *sermal* when spoken by the Iranian speaker.

At first glance, it would appear that this inhibitory effect contradicts the hypothesis of this study that experiential effects should lead to facilitation of variant forms of /θ/. However, I assert that it in fact supports the results showing facilitatory priming for Chinese participants hearing /s/ primes produced by the Chinese speaker. The /s/ primes produced by the Iranian speaker simply did not match closely enough the variant forms of /θ/ words in the mental lexicon of the Chinese participants in this study.

While this study certainly does not provide conclusive evidence that /θ/ substitutions do not function across speakers with different L1s, it does point to the need for future research. As mentioned above, the first step would be to perform acoustical analyses of the /s/ and /t/ substitutions produced by both speakers to determine where the differences lie. Using the speaker as a between-subjects factor, additional cross-modal priming experiments could then be conducted. These would help to determine whether the patterns of facilitation for both groups, and inhibition in the case of the Chinese, are the result of hearing two speakers, or whether they are linked solely to fine phonetic detail in the substitutions.

4.3.4 Question 4: Face Primes

Question number four in this study asked whether a subliminal face that matches the ethnicity and gender of the speaker would facilitate accented word recognition in the L2. The results were inconclusive because neither the ethnicity nor the gender of the face had any statistically significant effect on RTs or error rates in either the /θ/ or non-/θ/ items.

There are a number of possible reasons why the face primes failed to have any measurable effect. First, it may be that the display time of 50ms was too fast. Although I tested the masked prime images with two volunteers who were able to see the images at 66ms, they may simply be above average. It would have been preferable to test the images at each display time with separate volunteers to avoid the possibility that multiple exposure, in this case at 33ms and 50ms, is what led to recognition of the images at 66ms.

A second possible reason involves the degree to which participants may have developed a strategy of focusing their attention only on the target word because they learned that it would occur after the auditory prime word. Regarding semantic priming with orthographic words, the consensus has been that spreading activation is automatic, and it is not affected by attention or task context (Neely, 1991). However, recent research with orthographic primes has questioned this status quo. Naccache, Blandin, and Dehaene (2002) observed that in many masked priming experiments, participants learn to focus their attention on the window of time when the target appears. By manipulating the location of visual primes on the screen and the SOA from the prime offset to the appearance of the target, they demonstrated that priming effects disappear when participants' temporal attention is no longer focused to anticipate the appearance of the target. Although their findings may not apply to a masked image of a face as was used in this study, it is possible that the varying length of the auditory primes in my study (from 379ms to 1086ms) in some way blocked the visual face prime from having an effect. In other words, the lack of a regular time interval between offset of the face and onset of the target left participants "hanging", unsure of when the target would appear. This uncertainty then limited their attention to the visual prime image.

Using an alternative method to present the face primes may have improved the chances that they would have an effect on the processing of accented words in this study. One possibility would have been to force the attention of participants to the faces by simultaneously changing their location on the screen, while requiring that participants perform a variation on the dot-probe task. For example, Eberhardt and colleagues (2004) displayed masked face primes at one of four potential locations on the screen, and they required participants to indicate whether the "flash" had appeared on the left or the right side. In their between-subjects design, one group of participants saw 100 trials of Black faces, while the second and third groups saw a White face or no face respectively. The target images were then presented in the second half of the experiment, where response times were measured. This type of a between-subjects design, where half of the participants see a face that matches the foreign accent while the other half see a mismatching face, could potentially yield results in a study such as the one detailed in this thesis.

Chapter 5 Conclusion

This chapter includes a review of the aims of this study, followed by a summary of the results. I will also discuss possible implications of the results, limitations of the study, and suggestions for future research.

This study was designed to address the following research questions:

1) Does production of the English interdental fricative /θ/ correlate with performance on a lexical decision task that includes accented production of the same phoneme?
2) Does experience with the accent of English from one's home country facilitate word recognition of similarly-accented English?
3) Does it matter if the accent is imitated or authentic?
4) Does a subliminal face that matches the ethnicity and gender of the speaker facilitate accented word recognition in the L2?

I hypothesized that listeners would recognize /θ/ words more easily, and therefore more quickly and with fewer errors, when they were spoken with a substitution typical of their own accent, regardless of whether the accent was imitated or authentic. Words spoken with a /θ/ substitution that did not match the listener's, on the other hand, were expected to be more difficult to recognize and therefore elicit slower response times (RTs) and more errors. While all participants were hypothesized to react in this way, the effect was anticipated to be even clearer among participants who produce more accented tokens of /θ/ themselves. In addition, I predicted that a subliminal face that matched the accent and gender of the speaker of the prime word would facilitate recognition and thereby lead to faster RTs.

The study was carried out using two experiments. In Experiment 1, participants read aloud an English text, and their pronunciation of English /θ/ words was categorized to determine what substitutions, if any, they used. In Experiment 2, a cross-modal priming experiment, participants heard prime words containing /θ/ substitutions produced by two different speakers, and they then responded to a lexical decision task on canonical written versions of the same words. Together with two other native speakers of American English, I analyzed the audio files from Experiment 1. The RT results from Experiment 2 were analyzed in R (R Core Team, 2012) using linear mixed-effects models.

The results of Experiment 1 demonstrated that the documented /θ/ substitution preferences for Chinese and Persian speakers indeed applied to the participants in this study: the preferred substitution for /θ/ among the Chinese participants was /s/, while the Iranians more often used /t/. The data also showed that overall, the frequency of /θ/ substitution among the participants in this study was quite low: 22% for the Chinese and 36% for the Iranian participants. The results were inconclusive, however, in determining whether a correlation existed between the participants' production of /θ/ substitutes and their performance when hearing similar substitutions in the cross-modal priming experiment. Future research in this area requires a larger sample size, the use of the same words in both experiments, and possible collection of data from a non-experimental setting.

In Experiment 2, the results from the masked face prime were inconclusive because neither the ethnicity nor the gender of the face had any statistically significant effect on RTs or error rates in either the /θ/ or non-/θ/ items. It was suggested that the lack of an effect may have been due to the length of time the prime image was displayed, or the variation in the delay between the visual prime and target word.

The RT results from Experiment 2 showed that the participant groups varied in how they processed words depending on the prime condition (/s/, /t/, or control) and the accent of the speaker. The Chinese participants experienced significant facilitation when the /s/ prime was spoken by the Chinese speaker, and these results confirm the hypothesis of this study, suggesting that linguistic experience played a role in listening for those participants. Significant facilitation for the more-frequent variant forms, those with /s/ in this case, suggests that they are linked more strongly to the underlying lexical representations than the less-frequent /t/ variants (as in Ranbom & Connine, 2007). But while they showed facilitation for /s/ primes spoken by the Chinese speaker, the Chinese participants displayed an inhibitory effect for /s/ primes spoken by the Persian speaker, and no priming effects for the /t/ substitution.

The control experiment using non-/θ/ items helped to explain these results. When they heard non-/θ/ items that did not contain substitutions, the Chinese participants showed significant facilitatory priming effects regardless of the speaker of the prime. This suggests that it was not the Persian accent itself that led to the inhibitory effect, but rather the fact that the /s/ substitution occurred in an unexpected context. In other words, the mimicked nature of the /s/ prime from the Persian speaker disrupted processing in some way for the Chinese participants. I assert that these differing results for the Chinese participants, of facilitation and inhibition depending on the speaker, support the hypothesis that experience with substitution variants facilitates recognition. The /s/ primes produced by the Iranian speaker simply did not match closely enough the variant forms of /θ/ words in the mental lexicon of the Chinese participants in this study. Further research was suggested using acoustical analysis of the substitution variants produced by both speakers to determine how they differ in phonetic detail.

The Iranian participants, on the other hand, appear not to have been affected by the substitutions for /θ/ but rather by the accent itself, at least to a marginally significant degree. This pattern was confirmed by the results from the control experiment, where Iranian participants continued to show marginally significant facilitation when hearing their own accent, in this case for words with no segmental substitutions. These results were unexpected, but they coincide with previous findings showing that L2 speakers and listeners with the same L1 are better able to understand each other than those with different L1s (Bent & Bradlow, 2003).

These findings would appear to contradict those of Hanulíková and Weber (2011), where participants showed similar priming effects for variant forms regardless of the speaker of the prime. However, in their study, the accent of the prime was a between-subjects factor such that each group of participants only heard the variant forms produced by one speaker. Therefore, it is not possible to know whether the patterns observed for participants in this study are due to effects from the substitutions and accents themselves, or are the result of hearing two accents, one of which matched the participant's own. One possible way to explore this question in future research would be to use the accent of the prime as a between-subjects factor with some participants, and as a within-subjects factor with others.

The present study was limited to a relatively small sample size of 12 participants for each language group, making it difficult to generalize these findings to the larger population of L2 English learners. However, significant differences were observed in how each group of participants responded to segmental substitutions that were spoken by different speakers. This study provided empirical evidence that participants with different L1s do not necessarily process accented words produced by speakers with different L1s in the same way: segmental substitution may weigh more heavily for some groups, while others may be affected by phonetic detail of the accent as a whole. This constitutes a small, but important step towards a better understanding of how L2 listeners process foreign-accented words.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*(4), 419–439. doi:10.1006/jmla.1997.2558

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H. (2009). *languageR: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics."*

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion*, *15*(4), 501–520. doi:10.1080/02699930126251

Bates, D. M., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes*. Retrieved from http://CRAN.R-project.org/package=lme4

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600–1610. doi:10.1121/1.1603234

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press. Retrieved from http://ci.nii.ac.jp/naid/10018033931/

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775–794. doi:10.1121/1.1332378

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition*, *25*(4), 447–457. doi:10.3758/BF03201121

Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer*. Retrieved from retrieved

10 November 2012 from http://www.praat.org

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. doi:10.1016/j.cognition.2007.04.005

Bradlow, A. R., Pisoni, D., AkahansYamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English/r/and /l/: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, *101*(4), 2299–23.

Brannen, K. (2011). *The Perception and Production of Interdental Fricatives in Second Language Acquisition* (PhD Dissertation). McGill University, Montreal.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *The Journal of the Acoustical Society of America*, *117*(6), 3890–3901. doi:10.1121/1.1906060

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System*, *36*(1), 22–34. doi:10.1016/j.system.2007.11.003

Chang, J. (2001). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems*. Cambridge University Press.

Chen, T. H., & Massaro, D. W. (2004). Mandarin speech perception by ear and eye follows a universal principle. *Perception & Psychophysics*, *66*(5), 820–836. doi:10.3758/BF03194976

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335–359. doi:10.1016/S0022-5371(73)80014-3

Collins, A. F., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*, 193–210.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 113–121. doi:10.1037/0096-1523.14.1.113

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6),

3668. doi:10.1121/1.1810292

Dahan, D., & Magnuson, J. S. (2006). Spoken word recognition. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (pp. 249–284). London: Elsevier.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*(5-6), 507–534. doi:10.1080/01690960143000074

Davies, M. (2008). The corpus of contemporary American English: 425 million words, 1990-present. Retrieved from http://corpus.byu.edu/coca/

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, *87*(6), 876–893. doi:10.1037/0022-3514.87.6.876

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobstrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013–1027. doi:10.1037//0022-3514.69.6.1013

Fazio, R. H., Williams, C. J., & Powell, M. C. (2000). Measuring associative strength: Category-item associations and their activation from memory. *Political Psychology*, *21*(1), 7–25. doi:10.1111/0162-895X.00175

Fernandez, E. M., & Cairns, H. S. (2011). *Fundamentals of psycholinguistics*. Chichester: Wiley-Blackwell.

Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47–65.

Flege, J. E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, and Application* (pp. 565–604). Timonium, MD: York Press.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America*, *93*(3), 1589–1608. doi:10.1121/1.406818

Flege, J. E., McCutcheon, M. J., & Smith, S. C. (1987). The development of skill in producing word-final English stops. *The Journal of the Acoustical Society of America*, *82*(2), 433–447. doi:10.1121/1.395444

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language and Speech*, *38*(1), 25–55. doi:10.1177/002383099503800102

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, *41*(1), 78–104. doi:10.1006/jmla.1999.2638

Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and Foreign Accent Processing in English: Can Listeners Adapt? *Journal of Psycholinguistic Research*, *38*(4), 379–412. doi:http://dx.doi.org/10.1007/s10936-008-9097-8

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic bulletin & review*, *13*(3), 361–377.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*(5-6), 613–656. doi:10.1080/016909697386646

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1166–1183. doi:10.1037/0278-7393.22.5.1166

Graddol, D. (2000). *The future of English?* London: The British Council.

Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research*, *10*(3), 241–269. doi:10.1177/026765839401000304

Hanulíková, A., & Weber, A. (2010). Production of English interdental fricatives by Dutch, German, and English speakers. In *New Sounds* (pp. 173–178). Presented at the International Symposium on the Acquisition of Second Language Speech, Poznan, Poland.

Hanulíková, A., & Weber, A. (2011). Sink positive: Linguistic experience with th substitutions influences nonnative word recognition. *Attention, Perception, & Psychophysics*. doi:10.3758/s13414-011-0259-7

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. doi:10.1016/j.jml.2007.11.007

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for english as an international language. *Applied Linguistics*, *23*(1), 83 –103.

doi:10.1093/applin/23.1.83

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. doi:10.1016/0010-0285(89)90003-0

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008a). First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*, *19*(4), 332–338. doi:10.1111/j.1467-9280.2008.02090.x

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008b). First Impressions and Last Resorts How Listeners Adjust to Speaker Variability. *Psychological Science*, *19*(4), 332–338. doi:10.1111/j.1467-9280.2008.02090.x

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*(3), 245–294. doi:10.1016/0010-0277(91)90008-R

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods (Online)*, *44*(2), 325–343.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461. doi:10.1037/h0020279

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

Markham, D. (2007). Listeners and disguised voices: the imitation and perception of dialectal accent. *International Journal of Speech Language and the Law*, *6*(2), 290–299.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71. doi:10.1016/0010-0277(80)90015-3

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*(4), 653–675. doi:10.1037/0033-295X.101.4.653

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576–585. doi:10.1037/0096-1523.15.3.576

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. MIT Press.

Massaro, D. W. (2003). Read my tongue movements: Bimodal learning to perceive and produce non-native speech /r/ and /l/. In *Proceedings of Eurospeech (Interspeech)*. Presented at the 8th European Conference on Speech Communication and Technology, Geneva, Switzerland.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. doi:10.1016/0010-0285(86)90015-0

McQueen, J. M. (2005). Speech Perception. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of Cognition* (pp. 255–275). London: SAGE.

McQueen, J. M. (2007). Eight questions about spoken-word recognition. In M. G. Gaskell & G. Altmann (Eds.), *The Oxford Handbook of Psycholinguistics* (pp. 37–53). Oxford: Oxford University Press.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*(4), 323–331. doi:10.1016/0010-0277(79)90020-9

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*(2), 165–178. doi:10.1037/h0027366

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science*, *13*(5), 416–424. doi:10.1111/1467-9280.00474

Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: visual articulatory information enables the perception of second language sounds. *Psychological Research*, *71*(1), 4–12. doi:10.1007/s00426-005-0031-5

Neely, J. H. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 264–295). Routledge.

Neuhauser, S., & Simpson, A. P. (2007). Imitated or authentic? Listeners' judgements of foreign accents. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 1805–1808). Retrieved from http://www.researchgate.net

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234. doi:10.1016/0010-0277(94)90043-4

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395. doi:10.1037/0033-295X.115.2.357

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1209–1228. doi:10.1037/0278-7393.21.5.1209

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(03), 299–325. doi:10.1017/S0140525X00003241

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238. doi:10.1016/S0010-0285(03)00006-9

Paradigm 2.0. (2012). Perception Research Systems Incorporated.

Quillian, M. R. (1969). The teachable language comprehender: A simulation program and theory of language. *Communications of the ACM*, *12*(8), 459–476. doi:10.1145/363196.363214

R Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, *57*(2), 273–298. doi:10.1016/j.jml.2007.04.001

Rau, D. V., Chang, H.-H. A., & Tarone, E. E. (2009). Think or Sink: Chinese Learners' Acquisition of the English Voiceless Interdental Fricative. *Language Learning*, *59*(3), 581–621. doi:10.1111/j.1467-9922.2009.00518.x

Reis, M. S. (2006). *The perception and production of English interdental fricatives by Brazilian EFL learners* (unpublished Master's thesis). University of Santa Catarina, Brazil.

Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90*(1), 51–89. doi:10.1016/S0010-0277(03)00139-2

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology*, *5*(6), 546–554.

Soto-Faraco, S., Navarra, J., Weikum, W., Vouloumanos, A., Sebastián-Gallés, N., & Werker, J.

(2007). Discriminating languages by speech-reading. *Attention, Perception, & Psychophysics*, *69*(2), 218–231. doi:10.3758/BF03193744

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281–284. doi:10.1111/1467-9280.00151

Stilo, D. L., Talattof, K., & Clinton, J. W. (2005). *Modern Persian: spoken and written*. Yale University Press.

Strange, W. (1995). *Speech perception and linguistic experience: Issues in cross-language research*. York: York Press.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, *60*(4), 487–501. doi:10.1016/j.jml.2009.01.001

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*(6), 645–659. doi:10.1016/S0022-5371(79)90355-4

Tabossi, P., Burani, C., & Scott, D. (1995). Word identification in fluent speech. *Journal of Memory and Language*, *34*(4), 440–467. doi:10.1006/jmla.1995.1020

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. doi:10.1006/jmla.1998.2618

Vitevitch, M. S., Luce, P. A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: implications for the processing of spoken nonsense words. *Language and Speech*, *40*(1), 47–62. doi:10.1177/002383099704000103

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *68*(1–2), 306–311. doi:10.1006/brln.1999.2116

Wang, Y., Behne, D. M., & Jiang, H. (2008). Linguistic experience and audio-visual perception of non-native fricatives. *The Journal of the Acoustical Society of America*, *124*(3), 1716–1726. doi:10.1121/1.2956483

Weber, A., Broersma, M., & Aoyagi, M. (2011). Spoken-word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics*, *39*(4), 479–491. doi:10.1016/j.wocn.2010.12.004

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25. doi:10.1016/S0749-596X(03)00105-0

Wilson, L., & Wilson, M. (2001). Farsi speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems*. Cambridge University Press.

Winter, B. (2011). A very basic tutorial for performing linear mixed effects analyses. Retrieved from www.bodowinter.com

Appendices

Appendix 1: English Text Used in Experiment 1

Three siblings

Matthew, who had just celebrated his 50th birthday, had often undermined his parents' authority when he was younger. He was lazy and thick-headed. Once he was escorted home by the police for shoplifting sunglasses in one of the small shops in town! It was not until he was in his late twenties when, still at the threshold of becoming a criminal, it suddenly became clear to him that he wanted to be an actor. After numerous auditions at various theatres, he got the main part in "Big Thoughts of a Famous Thief". Three years later, he was even awarded a national prize for best performance in his role as a famous thief.

Emma was the only one of the three who had the ambition to obtain a degree of higher education. Her dream was to become a wealthy priest. Emma's parents had paid her tuition fees throughout the years, and it was rather unfortunate for them that Emma decided that becoming a priest had led her down the wrong path. She realized that she would better serve mankind as a soothsayer! Thanks to Emma's thorough work, she soon became the most respected soothsayer in town.

James, the youngest of the three children, had always been a worry to his parents because of his feeble health. There was always more than one disease that threatened James' health. Throughout his leisure time activities he had bruised his chest at least a thousand times and lost a few teeth. He seemed a lost cause and appeared like an actor who was clearly out of his depth. But things changed quickly after James' 30th birthday. Out of the blue, he decided to create a company on his own that would offer clients thrilling activities, including diving with sharks.

Appendix 2: English Self-assessment Questionnaire

Part A: Personal Information

Your field of study or job: ______________________________

Year of Birth: _________________

Gender: Female _____ Male_______

Country of Birth: _______________________________________

Part B: Language Background

Your First Language/Mother Tongue:________________________

Do you have a second Mother Tongue?:Yes_______ No________

If yes, what language? _________________________

English and other languages

In English, how do you rate your skills in each of these areas?

	Beginner	Intermediate	Advanced	Fluent
Reading				
Writing				
Speaking				
Listening				
Total				

Have you ever lived in an English-speaking country?

Yes_______ No________

If yes, where and for how many months/years?

____

____

Have you lived in another country/countries?

Yes _______ No _______

If yes, where and for how long?

____

____

____

Do you know any other languages in addition to your Mother Tongue/s and English?

Language	Level			
	Beginner	Inter.	Advanced	Fluent

<u>Part C: Other factors in language learning</u>

Do you have, or have you ever had vision problems (more than normal eyeglasses)?   Yes_______   No _________

____

Do you have, or have you ever had hearing problems?
Yes_______   No _________

____

Do you have, or have you ever had language-learning difficulties (e.g. SLI, reading/learning, etc.)?
Yes ______   No _________

____

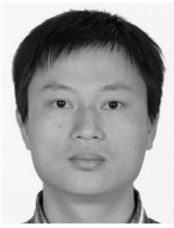
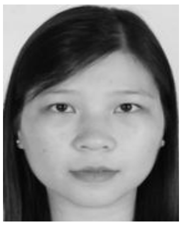
Have you been diagnosed with any other conditions that are thought to affect language learning? (e.g. ADHD, autism, etc.)?   Yes ______   No _________

____

Are you left-handed? Yes ______   No _________

Thank you for your participation!

Appendix 3: Post-test for Face Primes

Appendix 4: Experimental Stimulus Materials (12 /θ/ initial; 12 /θ/ final words)

Target	Accented primes		Control prime
	Chinese-accented	Persian-accented	
thesis /θiːsɪs/	/siːsɪs/	/tiːsɪs/	lurch
thief /θiːf/	/siːf/	/tiːf/	belly
throat /θroʊt/	/sroʊt/	/troʊt/	hero
theater /θiːətə(r)/	/siːətə(r)/	/tiːətə(r)/	pain
thirteen /θɜː(r)tiːn/	/sɜː(r)tiːn/	/tɜː(r)tiːn/	ink
thorough /θʌrəʊ/	/sʌrəʊ/	/tʌrəʊ/	carrot
theft /θɛft/	/sɛft/	/tɛft/	porter
thrive /θraɪv/	/sraɪv/	/traɪv/	merger
thirty /θɜː(r)tɪ/	/sɜː(r)tɪ/	/tɜː(r)tɪ/	glow
thermal /θɜrməl/	/sɜrməl/	/tɜrməl/	rhyme
thumb /θʌm/	/sʌm/	/tʌm/	apple
throw /θrəʊ/	/srəʊ/	/trəʊ/	news
fifth /fɪfθ/	/fɪfs/	/fɪft/	label
earth /ə(r)θ/	/ə(r)s/	/ə(r)t/	cook
south /saʊθ/	/saʊs/	/saʊt/	among
birth /bɜː(r)θ/	/bɜː(r)s/	/bɜː(r)t/	shell
beneath /bɪˈniːθ/	/bɪˈniːs/	/bɪˈniːt/	chicken
breath /brɛθ/	/brɛs/	/brɛt/	phrase
warmth /wɔː(r)mθ/	/wɔː(r)ms/	/wɔː(r)mt/	yell
health /hɛlθ/	/hɛls/	/hɛlt/	line
ninth /naɪnθ/	/naɪns/	/naɪnt/	mouse
wealth /wɛlθ/	/wɛls/	/wɛlt/	praise
length /lɛŋθ/	/lɛŋs/	/lɛŋt/	wooden
strength /strɛŋθ/	/strɛŋs/	/strɛŋt/	egg

Appendix 5: Control Experiment Stimulus Items (24 non-/θ/ words)

Target	Prime
port	tape
pale	career
boulder	ferry
bare	golden
zero	better
gap	movie
window	hair
cheek	lake
vote	forget
used	candy
ill	page
fame	pepper
fair	nose
rip	index
choke	valid
fear	paper
waiter	yeast
matter	book
zone	dense
cart	poke
church	seek
shame	chase
honey	bark
paste	junk

Appendix 6: Experiment 2 Instructions

Your task in this experiment is to decide, as quickly and as accurately as possible, whether a sequence of written letters is a real English word or not.

Each trial has four parts:

1. First, you will see a small circle in the center of the screen.
2. Then you will see a flash on the screen.
3. After that, you will hear a voice that will say something that may or may not be an English word.
4. Finally, you will see a sequence of letters.

If the sequence of written letters IS a real English word, press the GREEN button.

If the sequence of written letters is NOT a real English word, press the RED button.

Press any key on the button box to begin the practice session.
If you have any questions, please ask now.

The practice session is now finished.
When you press any key on the box, the experiment will begin.

If the sequence of written letters IS a real English word, press the GREEN button.

If the sequence of written letters is NOT a real English word, press the RED button.

Appendix 7: Fitted vs. Residual Values from the Model in Table 4.3