Auditory Priming Effects on the Production of Second Language Speech Sounds

by

Lindsay Ann Leong

B.A., University of British Columbia, 2009

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
Department of Linguistics
Faculty of Arts and Social Sciences

© Lindsay Ann Leong 2013

SIMON FRASER UNIVERSITY

Summer 2013

All rights reserved.
However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for "Fair Dealing." Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name:	**Lindsay Ann Leong**
Degree:	**Master of Arts (Linguistics)**
Title of Thesis:	***Auditory priming effects on the production of second language speech sounds***
Examining Committee:	**Chair:** Dr. Maite Taboada Associate Professor

Dr. Yue Wang
Senior Supervisor
Associate Professor

Dr. Trude Heift
Co-Supervisor
Professor

Dr. Christian Guilbault
External Examiner
Associate Professor
Department of French

Date Defended/Approved: August 13, 2013

Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website (www.lib.sfu.ca) at http://summit/sfu.ca and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, British Columbia, Canada

revised Fall 2011

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

Research shows that speech perception and production are connected, however, the extent to which auditory speech stimuli can affect second language production has been less thoroughly explored. The current study presents Mandarin learners of English with an English vowel as an auditory prime (/i/, /ɪ/, /u/, /ʊ/) followed by an English target word containing either a tensity congruent (e.g. prime: /i/ - target: “peach”) or incongruent (e.g. prime: /i/ - target: “pitch”) vowel. Pronunciation of the target word vowel following the two congruency conditions was measured in terms of vowel duration and formant frequency, as well as intelligibility and rating by native English listeners. Results showed that pronunciation of the front vowel contrast (/i-ɪ/) displayed more English-like formant frequency distribution and an increase in intelligibility in the congruent prime condition, suggesting that auditory speech information can positively affect the pronunciation of difficult second language speech contrasts.

Keywords: speech production and perception; priming; second language speech learning; Mandarin learners of English

To Dennis

Acknowledgements

I never would have even entered the program without Drs. Wang and Heift. Thank you both just for the opportunity.

I must also thank Dr. Yue Wang for her unwavering support throughout this process and for always being excited about my research. In over two years of work as a LABer she taught me the importance of taking notes, piloting experiments and being incredibly organized. I may not have always had the discipline to keep up with these things, but at least I can no longer deny their importance. I couldn't have asked for a better mentor. Thank you, Yue!

I would also like to thank Dr. Trude Heift for being the no-nonsense woman in my corner whenever I needed a pep talk and for the financial support early in my MA career. Thanks for all the fun chats in your office, Trude! I'm going to need your recipe for German cheesecake.

I will never be able to put into words the gratitude I have for all of my fellow LABers. I will miss you guys most of all, not just for your constant willingness to help me out but for your wonderful antics and friendship. I have to thank Melissa Philley for being my first ever research assistant and for always noticing when I had dropped a piece of my brain. Good luck at UBC! Alyssa Lee with her eager beaver attitude and unending smile made my life a million times easier and honestly, this thesis would not be finished right now without her invaluable assistance. Thank you to Parveen Kaila for always being on the ball and for always saying "Sure, no problem" when I had another pile of acoustic data for her. 3P0 forever, ladies. 3P0 forever.

Katelyn. You know what you mean to me. I'll never forget it.

A huge thank you to my fellow grad students for making Linguistics such an awesome department! Amber, Irina, Kayleigh – thanks for listening to me complain. I hope I returned the favour enough times to make it even.

Thank you to all of my research participants for taking the time out of your busy schedules to provide me with data and for being part of linguistic research.

My career as a MA student has been partially funded by Dr. Yue Wang and Dr. Trude Heift as a research assistant on some very exciting projects. Thank you also to the Dean of Graduate Studies and the Department of Linguistics at Simon Fraser University for the financial support.

Lastly, I get to thank my family. Dad for being my cheerleader. Steve for being impressed. Thomas for being unimpressed. Mom for the mental and emotional support, for never doubting me and for bragging about me to her friends. And Dennis for making sure I was clothed and fed and for watching me slowly go crazy and marrying me anyway.

Table of Contents

List of Tables

List of Figures

1. Introduction

While infants seem to be born with an innate ability to discriminate sound contrasts present in any human language, as the infant is exposed to the native language(s) in its environment, this ability begins to decline (Werker, Gilbert, Humphrey & Tees, 1981; Werker, 1989). Thus, adults who learn a second language (L2) generally have difficulty perceiving and producing the phones of the L2. One line of such L2 speech learning research has particularly focused on the perception and production link (Schneiderman, Bourdage & Champagne, 1988; Flege, 1991; Flege & Eefting, 1987; Flege, Bohn & Jang, 1997). For example, intensive perceptual training of L2 speech sounds can result in positive effects on the production domain (Bradlow, Akahane-Yamada, Pisoni & Tohkura, 1999; Bradlow, Pisoni, Akahane-Yamada & Tohkura, 1997; Akahane-Yamada, Tohkura, Bradlow & Pisoni, 1996; Rochet, 1995; Pimsleur, 1963). However, very few studies have examined how auditory primes specifically may affect the pronunciation of difficult L2 contrasts (Trofimovich 2005b; Trofimovich & Gatbonton, 2006) despite evidence that auditory primes are facilitative for native speakers (Jackson & Morton, 1984; Pilotti, Bergman, Gallo, Sommers & Roediger, 2000; Tilsen, 2009; Nielsen, 2007).

The aim of this thesis is to investigate the effects of single-segment auditory primes spoken by a native English speaker on the production of that segment for Mandarin learners of English. Mandarin learners of English were auditorily presented with a prime vowel and then asked to produce a target word containing the corresponding vowel. The stimuli included English vowel contrasts /i/-/ɪ/ and /u/-/ʊ/ of which the lax vowels /ɪ/ and /ʊ/ do not exist in Mandarin Chinese, while the tense vowels /i/ and /u/ are shared by both Mandarin and English vowel inventories (Wang, 1997). The results of this study will contribute to established literature on the connection between the production and perception of speech sounds for L2 learners and inform the cross-modal effects of auditory primes on pronunciation-focused targets.

1.1. Literature Review

Section 1.1 outlines a number of bodies of research regarding speech perception and production including theories of speech perception and second language speech learning in Section 1.1.1. The link between production and perception of speech will be discussed in Section 1.1.2 while specific studies using priming and shadowing techniques to study the perception/production link will be examined in Section 1.1.3. Section 1.1.4 outlines a summary of the research discussed in the literature review while Section 1.2 addresses the current study and its research questions.

1.1.1. Theoretical Framework

This study examines the link between perception and production with specific emphasis on how this link may be manifested within second language (L2) learners. Several theoretical perspectives offer insight for the current research including the Motor Theory (MT: Liberman & Mattingly, 1985) and direct realist (Fowler, 1986) approaches to speech perception which do not specifically address L2 learners but which offer predictions in terms of how perception and production are linked during speech. A direct realist approach is also embraced by the Perceptual Assimilation model (PAM: Best 1995; PAM-L2: Best & Tyler, 2007) that does offer predictions for L2 learners in terms of the perception and production of L2 sounds. Additionally, the Speech Learning model (SLM: Flege, 1995, 2007) focuses specifically on how L2 learners perceive and subsequently produce their L2.

Gestural theories of speech such as the Motor Theory of speech perception and the direct realist approach to speech perception postulate that the objects of perception are the gestural components of a speech act. MT argues that the objects of perception are the intended gestures (as opposed to actual vocal tract movements) made by the speaker during a given speech act and thus speech perception and production are "intimately linked" (Liberman & Mattingly, 1985, p. 2). From a direct realist perspective, the objects of perception are not intended gestures or motor commands but rather actual articulatory gestures performed by the speaker. These articulatory configurations are perceivable through the sound signal because the acoustic waveform is shaped by the gestures of the speaker and therefore carries crucial information about those gestures.

Thus both MT and direct realism argue for a perspective wherein speech perception and production are closely linked.

This link is also recognized within PAM-L2, which is based on the premises of a direct realist approach to speech perception. By embracing a gestural approach to perception, PAM-L2 necessitates a close relationship between perception and production wherein L2 learning in one modality is likely to be correlated with learning in the other. In terms of L2 learning, PAM-L2 suggests that new contrasts will be perceived in relation to existing contrasts and the level of difficulty experienced by the listener in creating a category for a new contrast is related to the perceived similarities and differences between the new contrast and the closest existing one(s). PAM-L2 makes a number of predictions in terms of how contrasts will be processed as a result of the relationship between the listener's L1 and L2. For instance, if two L2 phonological categories are assimilated into separate L1 categories, such a contrast should be relatively easy for the learner to discriminate. However, if two L2 phonological categories are perceived as equivalent to the same L1 phonological category, this would result in discriminatory difficulty and L2 listeners would need to create a new category in order to perceptually attune to this type of contrast.

The SLM also predicts that learners' L2 perception will be heavily influenced by existing L1 categories, however, SLM approaches the issue from a perceptual rather than a gestural perspective as embraced by PAM-L2 by making predictions as to how perception informs production learning. According to SLM, phonetic categories in the L1 and the L2 exist in a common perceptual space and those L2 categories which are perceived as more distant from existing L1 categories will be easier to discriminate than those which are perceived as perceptually similar. In terms of the connection between perception and production of L2 sounds, SLM predicts that the production of a sound will be learned as a result of establishing a corresponding phonetic category representation such that failure to perceptually discriminate L1 and L2 sounds will lead to production errors.

In summary, both SLM and PAM-L2 agree that linguistic experience can significantly affect how perception mediates articulatory learning. Based on the MT and direct realist predictions, PAM-L2 explains L2 speech learning as being a result of the listener perceiving articulatory gestures in the L2 speech signal and learning to detect higher-order linguistic information in order to facilitate more accurate pronunciation of difficult non-native speech sounds. The SLM predicts that L2 learners perceive acoustic features with references to the corresponding L1 sounds, which are then used to create new categories. The establishment of these phonetic categories is then hypothesized to facilitate accurate production of the newly acquired L2 sound

1.1.2. Production-Perception Link

Intuitively and from a layperson's perspective, it seems clear that the production and perception of speech should be linked as children and adults are able to both perceive and produce the sounds of their native tongues. From the perspective of gestural and motor theorists, such a connection exists to the extent that gestural information is perceivable through the sound signal and carries crucial information about the gestures that created the acoustic signal. According to the SLM and the PAM-L2, second language learners are expected to find certain L2 sounds to be difficult in terms of both perception and pronunciation, either because failure to discriminate perceptually has led to production errors (SLM) or because the gestural information within the L2 speech signal has been assimilated into a similar L1 contrast (PAM-L2). The research in the following section attempts to further illuminate the nature of the connection between perception and production with regards to both native and L2 speakers.

The extent to which motor circuits play a role during the perception of speech sounds has been extensively studied. For instance, transcranial magnetic stimulation of motor structures involved in certain speech sounds has been shown to both facilitate and disrupt perception during auditory presentation of the same sounds (D'Ausilio, Pulvermuller, Salmas, Bufalari, Begliomini, & Fadiga, 2009; Meister, Wilson, Deblieck, Wu & Iacoboni, 2007) suggesting that the motor cortex plays an important role in perception. In addition, a study by Fadiga, Craighero, Buccino and Rizzollati (2002) indicates that simply listening to speech sounds produces phoneme-specific activation of the motor areas involved in speech, findings which were replicated by Roy, Craighero, Fabbri-Destro and Fadiga (2008). Watkins and Paus (2004) observed increased excitability in motor areas correlated with activity in Broca's area during auditory speech perception and propose that activity in this area "primes" the motor system during perception, acting as a sort of interface between perception and articulation (p. 978). In fact, brain areas associated with speech gestures have been found to be activated during both audio and audio-visual speech perception by a number of researchers (e.g., Ojanen, Möttönen, Pekkola, Jääskeläinen, Joensuu, Autti, & Sams, 2005; Skipper, Nusbaum, & Small, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007; Wilson & Iacoboni, 2006; Wilson, Saygin, Sereno, & Iacoboni, 2004), findings which suggest a strong production-perception link neurolinguistically.

While neurolinguistic studies can give us a better understanding of how the brain uses and reacts to perceptual and motor commands related to speech use and comprehension, behavioural studies can tell us a great deal about how production and perception interact in real-world communication. Behavioural research by Cooper and Lauritsen (1974) has shown that after adaptation to a repeated auditory stimulus consisting of the syllable [pi], subsequent utterances of [pi] showed a significant decline in Voicing Onset Time (VOT) values as compared to repeated exposure to [i] indicating that perceptual adaptation had taken place. Additionally, Cooper, Blumstein and Nigro (1975) found similar results whereby repeated articulation of a segment facilitated the perception of subsequent speech stimuli. In a study by Elman (1981) the link between perception and production is examined by making real-time changes in auditory feedback to explore changes in production. Participants were asked to shadow (quickly reproduce) a number of auditory stimuli, some of which had been frequency-shifted in terms of pitch or fundamental frequency (F0). The results of this study showed that participants compensated their productions by adjusting their pitch up or down such that their utterances appeared more "normal" when compared to the frequency-shifted stimuli. Similar studies looking at fundamental frequency (Kawahara, 1998; Jones & Munhall, 2005) and formant frequencies (Houde & Jordan, 1998; Purcell & Munhall 2006; Villacorta, Perkell, & Guenther, 2007) during vowel production have shown that participants learn to adjust their production in order to compensate for alterations that change the vowel's perceived phonetic quality suggesting a strong link between perception and production during vocal processing. A study by Shiller, Sato, Gracco, and Baum (2009) found similar results and also showed that perceptual representations of the /s-ʃ/ contrast were changed and shifted in the same direction as feedback manipulation that perturbed the acoustic signal such that /s/ was closer in frequency to

/ʃ/. Accordingly, such adjustments to phoneme categories occur both articulatorily and perceptually. The results from each of these studies suggest that there exists a single mechanism that processes information during both speech perception and production.

In terms of L2 acquisition, the relationship between perception and production has been shown to be a positive correlational relationship wherein those sounds that are regularly mispronounced will also be misperceived and vice versa. For instance, Schneiderman et al. (1988) found that high perceptual ability by L2 learners as measured by a discrimination task correlated positively with more native-like pronunciation as judged by native listeners. Meanwhile, Flege (2003) stresses the importance of perceptual input in terms of production for L2 learners. In a series of related studies, he showed that Spanish learners of English who received input from primarily native English speakers produced more English-like VOTs for voiceless stops (Flege, 1991) versus Spanish learners who learned English from Spanish native speakers (Flege & Eefting, 1987). In a study of L2 learners of English from a variety of language backgrounds, Flege, Bohn and Jang (1997) were able to predict, in many cases, how well learners would perform in vowel perception tasks based on acoustic analysis of their productions of said vowels. That is, those learners with more native-like productions on certain vowels were better able to perceive those same vowels. However, not all of the variance in the production data was accounted for, and the researchers hypothesize that this could be because some aspects of perception change before production. They do not, however, rule out the possibility that production could change first, or could change without any corresponding changes in perception.

Additionally, a great deal of research has shown that training in perception of L2 speech can affect performance in production and vice-versa. For instance, perception-based training has been shown on numerous occasions to positively affect production in the L2 (Bradlow et al., 1999; Bradlow et al., 1997; Akahane-Yamada et al., 1996; Rochet, 1995; Pimsleur, 1963). In a perceptual training study for Japanese learners of English, it was found that the knowledge gained during perceptual learning of /r/ and /l/ was transferred to the production domain such that after training, the participants had improved pronunciation of these sounds without having received explicit training in production (Bradlow et al., 1997), a result which had long-term positive effects for production (Bradlow et al., 1999). Similar results were reported by Akahane-Yamada et al. (1996). Furthermore, Pimsleur (1963) conducted a training study in which American English learners of French received discrimination training in addition to laboratory practice (experimental group) or training in the form of laboratory practice only (control group). Training was followed by a pronunciation test wherein the discrimination training was found to be more effective in improving pronunciation, as judged by two native French listeners, than laboratory practice alone for French nasal vowels. Furthermore, Rochet (1995) found that perceptual training led to significant improvement for the perception of target stimuli as well as carryover to other syllables and was accompanied by improvement in production.

However, other studies into the connection between perception and production for L2 learners have yielded conflicting results. Wang (2002) for instance found that perceptual training on a subset of English vowel contrasts (/i/-/ɪ/, /u/-/ʊ/, and /ɛ/-/æ/) for native Mandarin and Cantonese learners produced significant perceptual improvements, but only marginal, non-significant improvements in production. Similar results were found in a training study on the English vowel pairs (/i/-/ɪ/, /u/-/ʊ/, and /ɑ/-/ʌ/) for Spanish-speaking learners of English (Garcia Perez, 2005). Moreover, Peperkamp and Bouchon (2011) found that performance on a perception task for French learners of English did not correlate with pronunciation in terms of accuracy for the English vowel contrast /i/-/ɪ/. This is hypothesized by the authors to be a result of their use of a speeded discrimination paradigm that could have interfered with the participants' ability to "subvocally rehearse the stimuli" (p. 164). Thus it seems that methodological differences could be, at least partially, responsible for some of the observed conflicting results.

Contributing to the level of misunderstanding surrounding this topic is the fact that very few studies have examined the effects of pronunciation training on the perception of L2 contrasts. In fact, some researchers have gone so far as to propose that accurate perception is necessary for pronunciation improvement (Henning 1966; Dreher & Larkins, 1972) or at least that perceptual mastery precedes native-like production (Flege & Eefting, 1987; Neufeld, 1980, 1988) thus suggesting that it is not possible for pronunciation training to affect perception. However, among the few studies that have evaluated pronunciation training techniques in terms of outcomes on production and perception, some results do suggest that articulatory training can affect perception. For instance, Perez-Gamboa (2001) was successful in showing improvement in both the production and perception of the English vowel contrast /i/ - /ɪ/ for Spanish speakers when pronunciation instruction was provided. Additionally, Leather (1990, as cited in Baese-Berk, 2010, p. 39) found that Dutch participants trained in Mandarin tone production were able to generalize this learning to perception with the caveat that only a single syllable was used during training and testing. On the other hand, a study looking at the perception and production of /r/ and /l/ by Japanese learners (Hattori, 2010 as cited in Baese-Berk, 2010, p. 39) showed that articulatory training, while effective for improving production, had no effect on perception.

Accordingly, previous work looking at the connection between the perception and production of speech has yielded mixed results. Some research has shown transfer or correlation between the two modalities (Schneiderman et al., 1988; Flege et al., 1997; Bradlow et al., 1999; Bradlow et al., 1997; Akahane-Yamada et al., 1996; Rochet, 1995; Pimsleur, 1963), while others have failed to show a robust link (Wang, 2002; Garcia Perez, 2005; Peperkamp & Bouchon, 2011) indicating that the relationship between production and perception is not straightforward and may be affected by such variables as L1 background for L2 learners, segment type (e.g. vowels versus consonants) and specific testing methods. However, understanding of this connection is very important to our overall understanding of language acquisition and thus additional research can only help to further illuminate the nature of this connection.

1.1.3. Priming and Shadowing

Priming tasks have been used to examine cross-modal interactions such as using visual primes and auditory targets (Holcomb & Anderson, 1993; Hernandez, Fennema-Notestine, Udell & Bates, 2001) to discover how semantic and lexical information is processed. Very few studies use auditory primes and spoken targets specifically and fewer still examine L2 learners and how they may respond to such tasks. However, such research is important for revealing how L2 production may be directly affected by auditory stimuli and such findings may inform further advancements in the perception-production link.

In the auditory-auditory priming domain, auditorily presented primes have been shown as having a facilitatory effect in terms of response time or accuracy for the recognition of repeated ("familiar") versus non-repeated ("novel") target stimuli and can be used to illustrate that auditory processing of words may be unconscious and unintentional (Trofimovich, 2005a). Consistently, Jackson and Morton (1984) found a large auditory priming effect on word recognition such that when participants were asked to identify words presented in noise, those words that had been previously presented auditorily to the participants were more readily recognized than control words. The results were not affected by the voice in which the prime words were spoken (male or female) and were much more robust than effects for visual primes. Similar findings were presented in a memory recall study by Pilotti, Bergman, Gallo, Sommers and Roediger (2000). Thus, it seems that auditory primes are facilitatory in terms of perception when auditory prime stimuli are presented some time prior to the testing phase and are therefore "familiar" when testing begins.

Despite the scarcity of cross-modal perception-production priming, studies on pronunciation using similar tasks such as shadowing and repetition may lend some references. Shadowing tasks involve a participant being asked to imitate the utterance of a model speaker immediately after presentation of the model stimuli. A vowel shadowing task by Tilsen (2009) showed that recently perceived vowels can influence the following pronunciation of those vowels in terms of sub-phonemic details, that is, subtle differences in vowel pronunciation that do not affect how the phoneme is categorically perceived by a listener. Similar results were found by Nielsen (2007) who showed that participants produced longer Voicing Onset Times (VOTs) after having been exposed to stimuli with prolonged VOTs. A study by Tilsen (2009) replicates the results from Nielsen (2007) and shows that speakers have internal representations of sub-phonemic details. However, participants did not regularly reduce their VOTs following shortened VOT primes suggesting that knowledge of linguistic contrast constraints places limitations on the extent to which a model can affect a speaker's pronunciation.

Using a repetition task, Trofimovich (2005b) predicted that auditory word priming effects would be involved in both L1 and L2 processing with a stronger effect in the participant's L1. In this study, participants listened to 80 words and were asked to repeat each word quickly and accurately after they heard each of prime stimuli. Response latency was measured and defined as the length of time between offset of the stimulus word and onset of the participant’s response. Results showed that in both the L1 and L2, word production initiation was faster following a repeated word than an unrepeated word suggesting that the same mechanism underlying auditory word priming supports the processing and learning of both a speakers’ first and second language. Trofimovich and Gatbonton (2006) further examined the effects of priming on L2 learners’ pronunciation, specifically in light of pedagogical issues within the realm of pronunciation teaching. They present experimental evidence suggesting that both repetition and focus on pronunciation can have benefits for L2 speech processing and suggest that L2 pronunciation teaching should include repetition tasks similar to auditory priming tasks as well as an emphasis on pronunciation. As mentioned, previously presented auditory stimuli appear to have facilitatory effects on perceptual tasks for native speakers and on perceptual as well as production tasks for native speakers and L2 speakers alike highlighting the similarities between L1 and L2 processing and learning.

The studies examined here seem to suggest that auditory primes can be an effective way to manipulate perceptual responses for native listeners (Jackson & Morton, 1984; Pilotti et al., 2000) as well as spoken responses for native listeners (Tilsen, 2009; Nielsen, 2007) and to some extent non-native listeners (Trofimovich 2005b; Trofimovich & Gatbonton, 2006). A number of variations on priming and shadowing paradigms have been examined here. For instance, in the Tilsen (2009) study, participants listened to a single vowel prime and were asked to imitate that vowel, while in the Nielsen (2007) study, participants were presented with an auditory word and then asked to produce that word. Trofimovich (2005b), Trofimovich and Gatbonton (2006), Jackson and Morton (1984) and Pilotti et al. (2000) presented their auditory priming stimuli in a study phase at some point prior to testing rather than prompting each response immediately following a single prime. Additionally, the latter two studies tested perception rather than production responses. However, research specifically involving single-segment auditory primes and spoken-word targets is less frequently found in the literature and the overall sparse nature of this type of work motivates the current study and in particular the specific priming paradigm chosen.

1.1.4. Summary

Research has shown that listeners seem to be able to perceptually extract articulatory information from the speech signal (Fowler, Brown, Sabadini & Weihing, 2003; Kozhevnikov & Chistovich, 1965; Porter & Castellanos, 1980; Porter & Lubker, 1980; Fadiga et al., 2002; Roy et al., 2008) and that L2 learners may even be able to use such information to more accurately produce the sounds of their L2 (Bradlow et al., 1999; Bradlow et al., 1997; Akahane-Yamada et al., 1996; Rochet, 1995; Pimsleur, 1963). These are important findings as adults learning an L2 often struggle with the phonetic inventory of their L2 (Flege, 1987; Flege, Munro & MacKay, 1995; Long, 1990; Oyama, 1976; Piske, MacKay & Flege, 2001), a phenomenon that has been a matter of interest for theorists such as Flege and Best when developing theories of L2 acquisition. Thus, more recent research has attempted to directly tap into the connection between perception and production by way of priming studies as a strategy to facilitate pronunciation for L2 learners (Trofimovich and Gatbonton, 2006).

1.2. The Current Study

The current study is motivated by the relatively small number of studies looking at the effect of auditory stimuli on spoken targets, specifically single-segment auditory primes and spoken-word targets and, particularly, with regards to L2 learners. As outlined in Section 1.1.3 some studies have shown that auditory stimuli can have facilitatory effects for L2 learners in terms of reaction time in response to primes (Trofimovich, 2005b) and pronunciation accuracy following repetition of auditory stimuli (Trofimovich & Gatbonton, 2006). However, little research has been done using auditory primes spoken by English speakers to examine how L2 learners of English produce the sounds of their second language. Thus, this study aims to contribute to established literature on the connection between the production and perception of speech sounds for L2 learners which has been previously concentrated on training studies (see Section 1.1.2) while expanding the priming paradigm to examine the cross-modal effects of auditory primes on pronunciation-focused targets.

In this study, Mandarin learners (as well as native English speakers as controls) were asked to complete a priming task wherein they were auditorily presented with a prime vowel and then asked to produce a target word containing the corresponding vowel. The stimuli included English vowel contrasts /i/-/ɪ/ and /u/-/ʊ/ of which the lax vowels /ɪ/ and /ʊ/ do not exist in Mandarin Chinese, while the tense vowels /i/ and /u/ are common in Mandarin and English (Wang, 1997). Stimuli were presented in either a congruent condition (prime vowel and target word vowel match in terms of tensity: e.g., /i/ - "peach") or an incongruent condition (prime vowel and target word are mismatched in terms of tensity: e.g., /i/ - "pitch"). Target word productions from the priming task were analyzed acoustically as well as evaluated by native English listeners for intelligibility and goodness rating. Additionally, in order to assess participant's baseline pronunciation and perception of the target vowels, prior to the main priming task, participants were also asked to complete (1) a speaking task to ensure that they displayed some measure of difficulty producing the two English lax vowels and to familiarize the participants with the target words, and (2) an auditory perceptual test in order to determine whether participants could discriminate between the four vowels.

These manipulations enable the examination of the primary research question of the current study: Does auditory priming of an English segment spoken by a native English speaker facilitate more accurate production of that segment for Mandarin learners of English? Additionally, are priming effects influenced by the proximity of a new L2 contrast (e.g., English /i/-/ɪ/) to existing L1 categories (e.g., Mandarin /i/) as predicted by models of second language learning such as the SLM and PAM-L2?

2. Methods

2.1. Overview

This is a priming study presenting native Mandarin learners of English and native English controls with a number of target words directly preceded by an auditory prime consisting of a single English vowel repeated three times. The participants were presented with the prime vowel triptych (e.g., /i,i,i/) and then asked to speak a target word (e.g., “peach”) into a microphone. Two congruency conditions were included. In the congruent condition the prime vowel triptych (e.g., /i,i,i/) matched the target word (e.g., “peach”) in terms of tensity. In the incongruent condition the prime vowel triptych (e.g., /i,i,i/) and the target word (e.g., “pitch”) did not match in terms of tensity. The stimuli for the priming portion of the experiment were prime vowels as spoken by native English speakers and written target words. Utterances of target words produced during the priming experiment were analyzed acoustically in terms of vowel duration and F1 and F2 frequency. Additionally, native English judges were asked to perform an intelligibility task by identifying the target word productions as well as perform a goodness rating task. Prior to the main priming task, the experiment also included an initial inclusion test as well as an auditory perceptual test.

2.2. Prime Stimuli

The prime stimuli chosen were four English vowels, /i/, /ɪ/, /u/, /ʊ/, of which the tense vowels, /i/ and /u/ are present in Mandarin whereas /ɪ/ and /ʊ/ are nonexistent. These target vowels were chosen because native Mandarin learners of English typically have difficulty pronouncing the lax vowels /ɪ/ and /ʊ/ in English words; however, the tense vowels /i/ and /u/ pose no problem for native Mandarin learners of English as they are both valid segments in Mandarin Chinese (Wang, 1997).

Articulatorily, the four target vowels (/i/, /ɪ/, /u/, /ʊ/) can be described as a high front tense (/i/) and lax vowel (/ɪ/) as well as a high back tense rounded (/u/) and lax rounded vowel (/ʊ/). According to Hillenbrand, Getty, Clark & Wheeler (1995), the average duration for these four vowels from longest to shortest is 275ms for /i/, 270ms for /u/, 221ms for /ʊ/ and 215ms for /ɪ/. Additionally, the average F1 and F2 frequencies for these four vowels can be plotted in the vowel space as displayed in Figure 1.

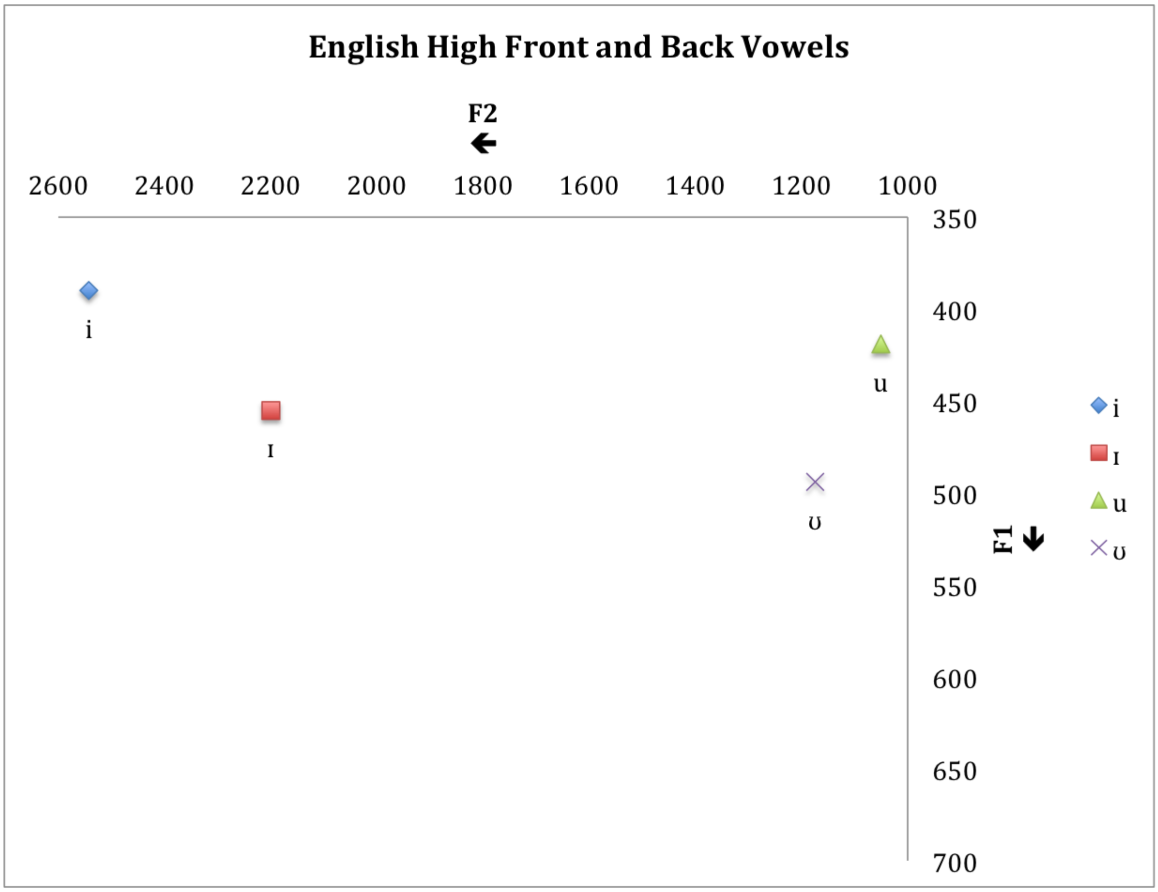

Figure 1. English high front and back vowels according to Hillenbrand et al. (1995)

All four target vowels were recorded by two native English speakers from Western Canada; one male and one female. The male speaker was 22 years old at the time of recording and was raised in Vancouver, British Columbia. The female speaker was 28 years old at the time of recording and was raised in Courtenay, British Columbia.

Each of the vowels was elicited by asking the speaker to read a single syllable word containing one of the four target vowels into a microphone. Recordings were made in a sound attenuated booth in the Language and Brain Lab at Simon Fraser University, using a high quality microphone (Shure KSM109 Condenser Microphone) and Audacity software. Care was taken to ensure that recording volume was appropriate for the individual speakers' normal speaking volume. All recordings for this experiment took place in the same location and with the same equipment.

The words containing the target vowels were *teak, tick, toque,* and *took*. Additionally, the words *tack, tuck and talk* were also recorded as distractors. These words were chosen by virtue of their being minimally different such that each vowel is preceded and followed by the same consonant in order to avoid colouring of the vowel by consonantal context.

Stimuli were presented using a MS PowerPoint presentation. Each slide contained the target word embedded in the sentence "Say ____ again". Presentation of the slides was randomized ahead of time using the online randomization tool http://www.random.org/lists/. Instructions presented at the beginning of the PowerPoint presentation explained that each slide would contain a phrase that the speaker would be asked to speak into the microphone, and to press a keyboard key in order to proceed to the next slide in the presentation. Speakers were instructed to repeat the entire phrase again if they felt their pronunciation was not ideal the first time and were reminded to take their time, as speed was not of importance for this task. Reminder slides displaying "Remember to read slowly" were repeated every seven slides in order to prevent the speakers from speeding up as the task went on.

2.2.1. Extraction of Prime Vowels

Once recording of the prime stimuli was complete, the recordings were subject to amplitude normalization using a script for the acoustic analysis software Praat, which analysed the sound files, calculated the optimal Root Mean Square (RMS) value and then normalized all files to that RMS value. Subsequent to normalization, the vowel targets were extracted from their consonantal contexts using Praat. The beginning and end of each vowel was measured using criteria laid out by Hirata and Whiton (2005) wherein the beginning of a vowel was "marked by the beginning of all formant frequencies and the beginning of a periodic waveform" and the end of the vowel was marked using “the termination of all visible formants” (p. 1649). A separate audio file was then created for each vowel token for ease of use in the priming task.

2.2.2. Prime Stimuli Evaluation

Each of the seven target and distractor words was recorded by the native English speakers a total of five times and the 70 extracted vowel recordings (7 vowels x 2 speakers x 5 repetitions) were then evaluated by three linguistically untrained, native English listeners (1 male and 2 females) who rated and identified each token vowel. The raters ranged in age from 21 to 32 years and each reported no history of speech, hearing or vision problems. All raters had some level of musical training experience, and all but the male participant reported some experience with a second language (Swedish, Spanish and Italian) at a moderate level or poorer.

The rating task was completed on a desktop computer using an experimental program developed with E-Prime. The experiment took place in a sound attenuated room within the Language and Brain Lab at Simon Fraser University using high quality headphones (AKG 141 Professional Studio Headphones).

Each rater was presented with the audio files containing the extracted vowel utterances one-by-one in random sequence. The seven target words (teak, tick, toque, took, tack, tuck, talk) were displayed on the screen with a number (1-7) displayed under each word. Listeners were instructed to listen to the vowel and then choose which word contained the vowel they heard by indicating on the computer’s numeric keypad the number under that word. Subsequently, the listeners were asked to rate the vowel on a scale from 1 (poor) to 5 (excellent). Any vowels that were incorrectly identified more than once were discarded as per Wang (1997) and any vowel tokens receiving a combined rating of less than 12 out of 15 were also discarded.

The best 14 recordings (one per vowel token for each of the two English speakers) as determined by the above criteria were chosen as the prime stimuli for the priming portion of the experiment.

2.2.3. *Target Stimuli*

Target stimuli for the priming task were 16 English words of relatively high frequency. The target words were arranged in minimal pair sets according to tensity such that for each tense/lax pair (/i/-/ɪ/ and /u/-/ʊ/) there were four minimal pair sets. Table 1 below shows each target word used in the priming task as well as its corresponding vowel tensity, frontness, IPA symbol and frequency within the Corpus of Contemporary American English (COCA) (Davies, 2008-).

Table 1. Target stimuli for priming task

Target Word	Vowel Tensity	Vowel Frontness	Vowel IPA symbol	Frequency per 450 million words (COCA)
Deed	Tense	Front	/i/	1915
Did	Lax	Front	/ɪ/	772,855
Feet	Tense	Front	/i/	92,284
Fit	Lax	Front	/ɪ/	33,999
Peach	Tense	Front	/i/	2825
Pitch	Lax	Front	/ɪ/	11,496
Sheep	Tense	Front	/i/	7160
Ship	Lax	Front	/ɪ/	28,725
Fool	Tense	Back	/u/	8348
Full	Lax	Back	/ʊ/	90,849
Luke	Tense	Back	/u/	7873
Look	Lax	Back	/ʊ/	289,258
Pool	Tense	Back	/u/	22,263
Pull	Lax	Back	/ʊ/	31,201
Who’d	Tense	Back	/u/	8945
Hood	Lax	Back	/ʊ/	7545

2.3. Priming Experiment

2.3.1. Overview

The priming experiment consisted of three parts. An inclusion test was given to ensure that Mandarin participants displayed some measure of difficulty producing the target English lax vowels. Additionally, the inclusion test served to familiarize the participants with the target words that would be used in the priming task. Following the inclusion test, the participants were given the priming task as well as an auditory perception test to determine whether participants could identify the target vowels.

2.3.2. Participants

Participants for the priming experiment were 20 (12 female and 8 male) learners of English whose first language is Mandarin Chinese as well as ten (5 male and 5 female) native English control participants. The Mandarin-speaking participants were mainly recruited from the student population at Simon Fraser University and ranged in age from 18 to 26. Others were recruited from the general population. In order to insure that the Mandarin-speaking participants were at an intermediate level in their English learning only participants who had lived in an English-speaking country for less than four years (Flege et al., 1997; Wang, Behne & Jiang, 2008) and began learning English in a classroom setting after the age of 10 (Moyer, 1999) were included. English participants ranged in age from 18 to 28 and were raised in the Western parts of Canada and the United States. Participants did not report any hearing, speech or neurological conditions and were asked to complete an Informed Consent form as well as a language background questionnaire before proceeding with the experiment.

2.3.3. Inclusion Test

The inclusion test was developed using E-Prime and presented the participant with a series of simple English sentences. Each sentence contained one of the 16 target words or one of six distractor words (Appendix A). The target words appeared in sentence-final position for consistency, with the exception of “who’d” which appeared in initial position as no grammatical English sentences end with such a construction.

Sentences were presented in random order and participants were instructed to press any keyboard key to advance to the next sentence. Five practice sentences were included at the beginning to give the participants a chance to get used to the procedures.

During the inclusion test equipment used and settings were identical to those used during the prime stimuli recording procedures described above. Experimenters read through the on-screen instructions with each participant, ensured that participants understood the procedures and clarified instructions if there was any confusion. Participants were instructed to speak in their typical speaking voice, and sound levels were adjusted to appropriately account for the speakers' natural speaking volume. Each participant was directed to read each sentence displayed on the computer screen in their normal speaking voice into the microphone and to press any keyboard key to prompt the computer to display the next sentence.

Scoring was completed by transcription of the target word into IPA by the principal investigator and a second experienced transcriber. A Mandarin speakers' recordings were included if the speaker received a score of less than 50% for all target words. Utterances wherein the target vowel did not match the expected or native English pronunciation (/i/, /ɪ/, /u/, /ʊ/) were scored as incorrect. In total, utterances from 12 female and 8 male Mandarin speakers were included as well as all English-speaking control participants.

2.3.4. Priming Task

Immediately after the inclusion test, participants were asked to complete the priming task. The procedures for the priming task were also developed using E-Prime. For each trial the priming vowel was repeated 3 times followed immediately by presentation of the target word on the computer monitor in either the congruent condition (prime vowel and target word vowel match for tensity) or incongruent condition (prime vowel and target word do not match for tensity). Figure 2 shows a visual representation of an example priming task trial.

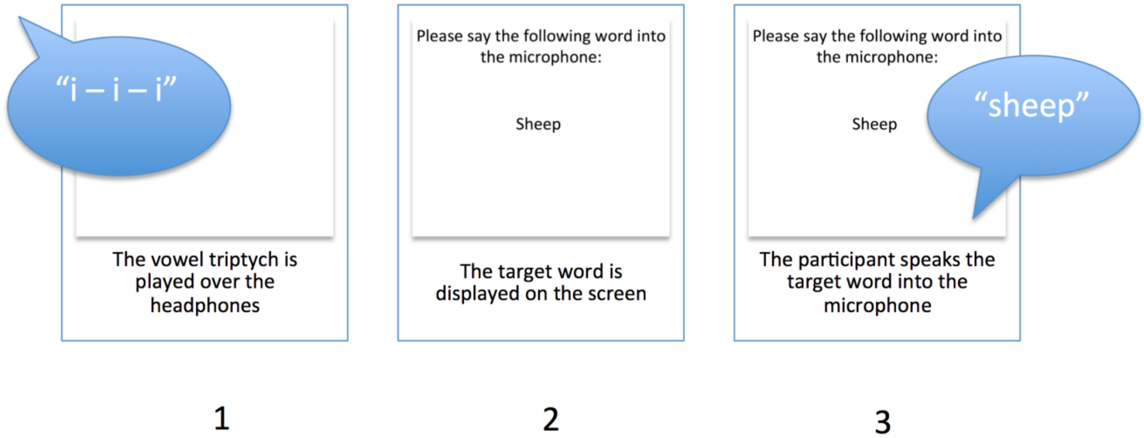

Figure 2. Example priming task trial with the high front tense vowel in a congruent condition

During the priming task, participants listened to the prime vowels over the headphones and immediately following the prime the computer monitor would display one of the 16 target words (see Table 1) which the participant would then read into the microphone. Interstimulus intervals of 250ms occurred between each of the three repetitions of the prime vowel as well as between the final prime vowel and presentation of the target word (Spencer & Wiley, 2008). Participants were reminded to speak each target word using their normal speaking voice as soon as the word was displayed and to press any keyboard key to advance to the next trial making sure to be careful not to cut off their recording by advancing too quickly. Each participant was instructed to focus on the target word so as to avoid confusion such as repeating the prime vowel rather than reading the word as displayed on the screen. Participants were also instructed to move onto the next trial should they make a mistake during pronunciation of the target word. There were a total of 64 priming task trials (16 words x 2 speakers x 2 congruency conditions) as well as 16 practice trials that did not use any of the four target vowels or words.

2.3.5. Perceptual Test

A separate perceptual test was administered after the priming task. During the perceptual test, a computer monitor displayed a set of minimal pairs (e.g., peach, pitch) under which are the numbers 1 and 2, while concurrently playing one of the two words in the minimal pair set over the headphones. Participants were instructed to use the numeric keypad to indicate which of the two words they heard over the headphones.

Stimuli for the perceptual test were taken from the original prime vowel recordings before extraction of the target vowel and were not used as target words in the priming task. These words were *teak, tick, took, toque, tack, tuck* and *talk*. Table 2 below shows the minimal pairs used in the task. Display of minimal pair sets was counter balanced such that the correct answer for each word was “1” for half of the trials and “2” for the remaining half of the trials. Each word was presented a total of four times (1 female production, 1 male production repeated twice each) for a total of 28 trials. Trial presentation was randomized.

Table 2. Target and distractor minimal pairs for the perceptual test.

Minimal pairs including target vowels	Minimal pairs including distractor vowels
teak – tick	talk – tack
tick – teak	tack – tuck
took – toque	tuck – talk
toque – took	

Before commencing with the experiment, participants were asked to inform the experimenter if they were unfamiliar with the pronunciation of any of the target words. If the participant indicated that they did not recognize a word, the experimenter would pronounce the word as well as offer a definition.

2.4. Acoustic Analysis

Acoustic analyses of vowel duration as well as formant frequency (F1 and F2) measurements were completed. According to a study by Kondaurova and Francis (2008) native English listeners rely predominantly on formant frequencies and duration to distinguish English tense and lax vowels. In terms of articulatory-acoustic correspondence, F1 corresponds to the height of the tongue during vowel production while F2 corresponds to tongue backness (e.g., Rogers, 2000). Therefore these acoustic measurements will provide relevant phonetic information for how each vowel is pronounced.

2.4.1. Duration

Vowel duration was measured for all target words spoken by both the Mandarin learners of English and the native English control participants. The beginning and end of each vowel was measured using the same criteria utilized for creation of the prime stimuli wherein the beginning of a vowel was marked by the start of all formant frequencies and periodic wave form and the end was marked by the end of any visible formants (Hirata & Whiton, 2005). Praat was used for all duration measurements. Figure 3 shows a screenshot of Praat and where the beginning and end of the vowel would have been measured.

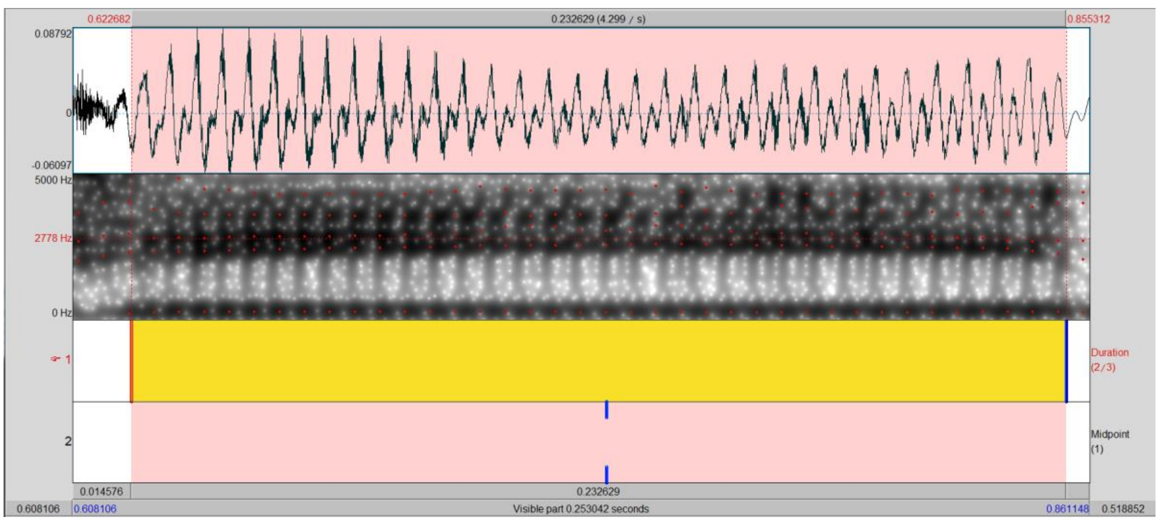

Figure 3. Screenshot of Praat showing how duration and midpoint were measured for the word "deed"

Each set of duration data was double checked by an additional researcher who re-measured two randomly selected tokens representative of each of the four target vowels. Thus, 8 total tokens per participant were re-measured. Notes were made wherever the two measurements differed by more than 10% as per the just noticeable difference. The principal investigator subsequently reviewed all double-checked measurements and made any final adjustments.

2.4.2. Formant Frequency

Formant frequency was measured automatically using the formant detection feature of Praat software. Both F1 and F2 were measured for all tokens spoken by both the English and Mandarin participants at the midpoint of each vowel in order to avoid colouring of the vowel by surrounding consonants. Each measurement was reviewed manually using Hillenbrand et al. (1995) as a reference point to ensure accuracy as the formant detection feature of Praat is often unreliable in more ambiguous cases (e.g., where F1 and F2 are close together).

2.5. Native English Judgement

2.5.1. Overview

Following completion of data collection for the priming experiment, 16 (11 female and 5 male) Native speakers of English were recruited to identify and judge the utterances. Judges ranged in age from 18-49 years of age and were raised in Western Canada and United States. None of the judges reported having learned a second language beyond a moderate level and none reported any hearing, speech or neurological conditions.

2.5.2. Stimuli

The stimuli used for the native English judgement task were those utterances that were elicited from the native Mandarin learners of English as well as the native English controls during the priming task.

Among all the productions, inappropriate utterances were excluded. These included such phenomena as incomplete tokens where the speaker may have cut off their recording by pressing a keyboard key too soon and coughing or other non-speech vocalizations. Recordings that included excessive dead space before and after the target word or audible keyboard noise were trimmed to eliminate these distractions. Finally, the recordings were subject to the same amplitude normalization script as was used in the creation of the prime stimuli in order to insure that all stimuli were equally audible for the listener.

Stimuli were organized into five separate versions of the experiment such that each version included the utterances of four Mandarin-speaking and two English-speaking participants. Each of the five versions of the experiment was given to at least three English-speaking judges. Thus each speaker would have their utterances evaluated by at least three judges, and each judge would evaluate six speakers. Table 3 shows all speakers and judges and how they were systematized for the judgement task.

Table 3. All speakers and judges

Judge	Mandarin Speaker	English Speaker
1 2 3 4	1 2 3 4	1 2
5 6 7	5 6 7 8	3 4
8 9 10	9 10 11 12	5 6
11 12 13	13 14 15 16	7 8
14 15 16	17 18 19 20	9 10

2.5.3. Inclusion Test

In order to insure inclusion of only those native English judges who were capable of discriminating between the four token vowels, each judge completed an inclusion test to determine eligibility. The inclusion test in this case was identical to the perceptual test described in Section 2.3.5. Evaluation of each judge's scores was completed on the spot and all received a score of 100%.

2.5.4. Intelligibility and Rating

Following successful completion of the inclusion test, judges were asked to complete the judgement task. Stimuli for this task included the elicited utterances from the priming task. Procedures for the intelligibility task were adapted from the perceptual test and instructions, counter-balancing, randomization and response mechanisms were identical. Presentation of stimuli was mixed across Mandarin and English speakers. A computer monitor displayed a set of minimal pairs (e.g., hood, who'd) under which were displayed the numbers 1 and 2, while concurrently playing one of the two words in the minimal pair set over the headphones. Judges were instructed to use the numeric keypad to indicate which of the two words they heard over the headphones. Response with the keyboard prompted the program to display instructions to rate only the vowel in the word on a scale from 1 (poorly pronounced) to 5 (excellently pronounced). Response on the rating portion of the task prompted the program to move onto subsequent trials.

3. Results

Measurements for each dependent variable set (duration, F1, F2, intelligibility and goodness rating) were each analyzed using repeated measures ANOVAs with tensity (tense vowels and lax vowels) and congruency (congruent and incongruent prime and target vowel) as within-group factors and group (English and Mandarin speakers) as the between-subjects factor. Additional ANOVAs were run where necessary.

3.1. Acoustic Analyses

3.1.1. Duration

Table 4 represents mean vowel duration in milliseconds (ms) for the tense and lax front and back vowels for English and Mandarin speakers in each of the two congruency conditions. For front vowels, significant main effects were found for tensity [$F(1,28)=41.88$, $p<.001$] where, as expected, tense vowels (167 ms) were shown to be longer than lax vowels (126 ms). Similar significant tensity effects were observed for back vowels [$F(1,28)=49.99$, $p<.001$], again with tense vowels (183 ms) being longer than lax vowels (161 ms). Likewise, for back vowels, a significant effect of group was exhibited [$F(1.28)=9.32$, $p=.005$], with Mandarin speakers producing longer vowels (198 ms) than English speakers (147 ms). Additionally, a significant interaction for tensity x group [$F(2,28)=30.28$, $p<.001$] was found for back vowels. No other significant main effects or interactions were found.

A follow-up one-way ANOVA was run in order to examine the effect of tensity for the two groups as motivated by the significant tensity x group interaction for the back vowels. Results showed that English speakers distinguished tense and lax back vowels in both congruent [$F(1,9)=23.4$, $p=.001$] and incongruent [$F(1,9)=31.8$, $p<.001$] conditions. In the congruent condition the tense vowel was longer than the lax vowel (tense: 169 ms; lax: 130 ms). This was also the case in the incongruent condition (tense: 164 ms; lax: 123 ms). However, Mandarin speakers did not distinguish tense and lax back vowels in terms of duration in either of the congruency conditions.

Table 4. Mean vowel duration (in ms) for high front and back, tense and lax vowels produced by English and Mandarin speakers in the congruent and incongruent conditions. Standard deviations are given in parentheses.

	English		Mandarin	
	Congruent	**Incongruent**	**Congruent**	**Incongruent**
/i/	151 (24)	150 (24)	184 (66)	182 (61)
/ɪ/	115 (26)	111 (24)	138 (44)	141 (46)
/u/	169 (36)	164 (31)	204 (51)	196 (51)
/ʊ/	130 (20)	123 (21)	199 (51)	191 (55)

3.1.2. Formants

For front vowels, F1 analysis showed significant main effects for tensity [$F(1,28)=109.14$, $p<.001$] where the tense vowel /i/ (364Hz) was lower than the lax vowel /ɪ/ (446 Hz). Additionally, significant interactions were found for tensity x group [$F(1,28)=57.71$, $p<.001$] and tensity x group x congruency [$F(1,28)=6.51$, $p=.016$]. F2 analysis also showed significant main effects for tensity [$F(1,28)=273.3$, $p<.001$] with the tense vowel /i/ (2421 Hz) being higher than the lax vowel /ɪ/ (2174 Hz). Additionally, significant interactions were found for tensity x group [$F(1,28)=153.7$, $p<.001$] and tensity x group x congruency [$F(1,28)=6.13$, $p=.020$].

For back vowel F1 analysis, significant main effects were found for tensity [$F(1,28)=171.56$, $p<.001$] with the tense vowel /u/ having a lower F1 (410 Hz) than the lax vowel /ʊ/ (472 ms). Additionally, a significant interaction for tensity x group [$F(1,28)=118.54$, $p<.001$] was found. No other significant main effects or interactions were found for F1, nor were there any significant main effects or interactions for F2 in the back vowel contrast.

Figure 4 below shows average F1 and F2 values (in Hz) for both Mandarin and English speakers as plotted on a vowel chart. Values for congruent and incongruent conditions are pooled.

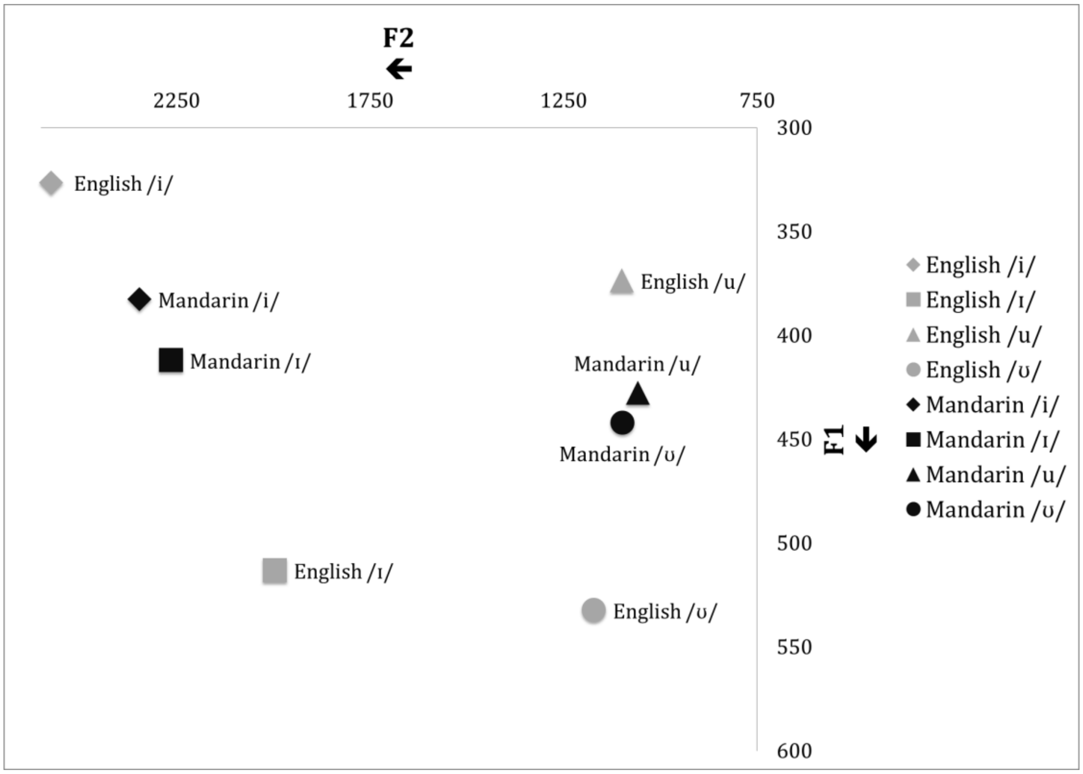

Figure 4. ***Average F1 and F2 values (in Hz) for the high front and back tense and lax vowels produced by English and Mandarin speakers plotted in a vowel chart***

Returning now to front vowel analysis of F1, two-way ANOVAs for tensity and congruency were motivated by the above-mentioned significant three-way interaction. A significant main effect of tensity was found for both English [$F(1,9)=89.5$, $p<.001$] and Mandarin [$F(1,19)=7.3$, $p=.014$] groups. Additionally, a significant interaction of tensity and congruency was found for Mandarin participants [$F(1,19)=13.3$, $p=.002$]. Based on this interaction, one-way ANOVAs were run for each tensity condition with congruency as a factor for the Mandarin speaking group. Significant congruency effects were found for both the tense [$F(1,19)=9.48$, $p<.05$] and lax [$F(1,19)=8.03$, $p<.05$] vowels. Mean F1 for the tense vowel was lower (approximating the native direction) in the congruent (376 Hz) versus incongruent (388 Hz) condition. Similarly for the lax vowel, mean F1 was higher (approximating the native direction) in the congruent (418 Hz) versus incongruent (405 Hz) condition. Thus, these results indicate that congruent prime vowels lead to more native-like patterns for the F1 of high, front tense and lax vowels as produced by Mandarin learners of English.

The tests described above were also run on the F2 data for front vowels based on the corresponding significant 3-way interaction. Two-way ANOVAs for tensity and congruency with each language group showed a main effect of tensity for English participants ([$F(1,9)=201$, $p<.001$]) and no other significant results. While Mandarin participants also showed a main effect of tensity ([$F(1,19)=17.5$, $p=.001$]), an interaction of tensity and congruency was also found ([$F(1,19)=12.4$, $p=.002$]). Further one-way ANOVAs with congruency as a factor showed a significant effect of congruency for the lax vowel for the Mandarin speakers ([$F(1,19)=23.4$, $p<.001$]), with the lax vowel F2 in the congruent condition (2238 Hz) being lower (approximating the native direction) than in the incongruent condition (2287 Hz). Additionally, a further one-way ANOVA with tensity as a factor was examined showing that for Mandarin speakers in the congruent condition, tensity was significant ([$F(1,19)=23.2$, $p<.001$]) wherein the tense vowel was higher (2363 Hz) than the lax vowel (2238 Hz). However, in the incongruent condition Mandarin speakers were unable to distinguish tense and lax vowels with regards to F2. As for the F1 data, results indicate that congruent prime vowels lead to more native-like patterns overall for Mandarin speakers producing front tense and lax vowels.

Based on the significant statistical differences as a function of congruency, Figure 5 displays the high front tense and lax vowels spoken by Mandarin participants in both the congruent and incongruent conditions along with English tokens (pooled across congruency conditions as no prime effect was found) as reference points. The figure illustrates the movement of the tokens produced by Mandarin speakers in a more native-like direction in the congruent condition as compared to the incongruent condition.

Figure 5. ***High front tense and lax vowels as spoken by Mandarin learners of English in congruent and incongruent prime conditions. English tokens (pooled across congruency conditions) are included as reference points***

In sum, formant frequency analyses revealed that for both tense and lax front vowels, congruency had a significant effect in terms of how Mandarin speakers produced F1 and F2. Specifically, productions following a congruent prime segment were pushed in the direction of more native-like formant frequencies. However, this was not the case for back vowels and neither were there any significant effects of congruency for English speakers.

3.2. Native Listener Judgements

3.2.1. Intelligibility

The intelligibility of the target vowels produced by English and Mandarin speakers were evaluated by native English listeners in an identification task. Table 5 represents mean percent correct totals for the tense and lax high front and back vowels for English and Mandarin speakers in each of the two congruency conditions.

Table 5. Mean percent correct for high front and back, tense and lax vowels produced by English and Mandarin speakers in the congruent and incongruent conditions as judged by native listeners of English. Standard deviations are given in parentheses

	English		**Mandarin**	
	Congruent	**Incongruent**	**Congruent**	**Incongruent**
/i/	96 (3)	96 (7)	65 (23)	57 (26)
/ɪ/	100 (1)	99 (2)	65 (21)	56 (24)
/u/	96 (4)	95 (5)	50 (21)	45 (25)
/ʊ/	95 (11)	97 (8)	56 (21)	54 (21)

For the front vowels, a three-way ANOVA revealed significant main effects for congruency ([$F(1,28)=8.61$, $p=.007$]) with utterances in the congruent condition being correctly identified 76% of the time and utterances in the incongruent condition 70% of the time. There was also a significant effect of group [$F(1,28)=156.98$, $p<.001$] where utterances produced by English speakers were correctly identified 97% of the time and utterances by Mandarin speakers were correctly identified 61% of the time. Additionally, a significant interaction for congruency x group [$F(1,28)=6.81$, $p=.014$] was revealed. No other significant main effects or interactions were found.

Significant interactions of congruency and group motivate further one-way ANOVAs for front vowels. Effects of congruency showed that for both tense and lax vowels, Mandarin speakers showed a significant effect of congruency (tense: [$F(1,19)=7.3$, $p=.014$]; lax: [$F(1,19)=8.8$, $p=.008$]) while English speakers did not. In this case, both tense and lax vowels in the congruent condition were identified more accurately (Tense: 65%; Lax: 65%) than those in the incongruent condition (Tense: 57%; Lax: 56%).

For back vowels, significant main effects were found for group [$F(1,28)=435.74$, $p<.001$] where utterances by English participants were more accurately identified (96%)

than those by Mandarin speakers (52%). No other significant main effects or interactions were found.

3.2.2. Rating

Table 6 represents mean rating totals (on a scale of 1 to 5) for the tense and lax high front and back vowels for English and Mandarin speakers in each of the two congruency conditions. For front vowels, significant main effects were found for group [$F(1,28)=213.08$, $p<.001$] where English speakers received higher ratings (4.8) than Mandarin speakers (3.8) overall. Consistently, for back vowels, significant main effects were found for group [$F(1,28)=178.73$, $p<.001$] where utterances from English speakers were more highly rated (4.7) than those from Mandarin speakers (3.0). No other significant main effects or interactions were found.

Table 6. Mean ratings (on a scale of 1 to 5) for high front and back, tense and lax vowels produced by English and Mandarin speakers in the congruent and incongruent conditions. Standard deviations are given in parentheses.

	English		Mandarin	
	Congruent	Incongruent	Congruent	Incongruent
/i/	4.8 (0.2)	4.7 (0.2)	3.3 (0.4)	3.2 (0.4)
/ɪ/	4.9 (0.1)	4.8 (0.1)	3.4 (0.3)	3.3 (0.5)
/u/	4.7 (0.2)	4.7 (0.2)	3.1 (0.5)	3.0 (0.7)
/ʊ/	4.7 (0.1)	4.7 (0.2)	3.0 (0.5)	2.9 (0.5)

3.3. Perceptual Test Analysis

The purpose of the perceptual test, which was administered separately after the speaking task, was to determine whether Mandarin participants could adequately distinguish between the two vowel pairs. Repeated measures ANOVAs were run on the perceptual test data with tensity (tense vowels and lax vowels) and frontness (front vowels and back vowels) as within-group factors and group (English and Mandarin) as the between-subjects factor.

Table 7 shows the mean percent correct for each of the four vowels on the perceptual test for both English and Mandarin speakers. Results of a three-way ANOVA show a main effect of group ([$F(1,28)=13.3$, $p=.001$]) which indicates that English participants performed better (88%) than Mandarin participants (73%) overall. No other significant effects or interactions were found. Although Mandarin participants did performed significantly poorer than the English participants overall, they did perform better than chance for all four vowels indicating that the Mandarin speakers were able to perceive tense and lax vowels.

Table 7. Mean percent correct for high front and back, tense and lax vowels on the perceptual test for English and Mandarin speakers. Standard deviations are given in parentheses.

	English	**Mandarin**
/i/	90 (13)	70 (24)
/ɪ/	83 (21)	74 (25)
/u/	90 (13)	76 (19)
/ʊ/	90 (13)	70 (19)

3.4. Summary of Results

Results revealed that congruent prime segments did seem to facilitate more native-like formant frequencies for the Mandarin speakers than incongruent prime segments for front vowels. On the other hand, formant frequency analysis showed an overall lack of a congruency effect for back vowels for the Mandarin speakers and congruency of the prime segment did not seem to have an effect on the duration of the vowel pairs for Mandarin speakers who seemed able to distinguish vowel duration at least as well as the English participants across congruency conditions.

As was the case for formant frequency, there was an effect of congruency for the intelligibility of Mandarin productions for the tense and lax front vowels with those in the congruent condition being better identified by native English judges than those in the incongruent condition. This could indicate that the movement of Mandarin-produced vowels through the vowel space towards more native-like tokens in the congruent conditions did result in increased intelligibility. However, congruency was not a significant factor for the intelligibility of Mandarin-produced back vowels and only a significant effect of group was found for rating data indicating that tokens produced by English speakers were more highly rated than those by Mandarin speakers. Lastly, acoustic analyses as well as intelligibility and rating analyses showed no significant effect of congruency for English speakers, likely due to a ceiling effect.

Results from the perceptual test showed that while Mandarin speakers were overall less proficient at vowel identification than the English speakers, they did perform better than chance for all four vowels indicating that the Mandarin speakers were able to perceive tense and lax vowels.

4. Discussion

The discussion section is divided into four main subsections. Section 4.1 will discuss the results of the acoustic analyses for duration (Section 4.1.1) and formant frequency (Section 4.1.2) while Section 4.2 will discuss the results of the native listener judgements. A general discussion of the findings is found in Section 4.3 as well as how the current results integrate into the previous research. Lastly, Section 4.4 will summarize and relate the current findings to the theoretical models discussed in Section 1.1.1.

4.1. Acoustic Analysis

4.1.1. Duration

Results from the durational analysis showed that Mandarin speakers patterned with native English speakers in producing the tense vowel /i/ with longer durations than the lax vowel /ɪ/. This is expected as previous research has revealed that Mandarin speakers tend to exaggerate the durational differences between the tense and lax pair /i/ and /ɪ/ (Wang 1997). This is also reflected in the perceptual domain as Bohn (1995) showed that native Mandarin speakers rely exclusively on durational cues for differentiating the English vowel pair /i/-/ɪ/. The tendency for Mandarin speakers to attend to durational cues for this vowel pair is also reflected in the results of the current study in terms of the relative lack of effect of congruent prime segments on duration for Mandarin speakers' productions. That is, Mandarin speakers seem to already be using duration to distinguish tense and lax vowels, thus prime segments unsurprisingly have little effect on this particular phonetic contrast.

Mandarin speakers were still unaffected by priming effects for the back vowel contrast, but in this case they did not distinguish between the tense and lax vowels at all. Again, this was shown to be the case previously where Mandarin productions of /ʊ/ were not significantly shorter than the tense vowel /u/ (Wang 1997). However, this becomes more difficult to explain in terms of priming effects as one might expect a congruent prime to then exact a facilitatory effect in this case where Mandarin speakers are not already attending to duration and, especially, in light of the fact that in other situations Mandarin speakers exclusively use durational cues as in for the front vowel contrast.

4.1.2. Formants

F1 and F2 analyses for the front vowels appear to display priming effects for the Mandarin speakers. Mean F1 for the tense vowel was lower and F2, while not significant, was higher in the congruent condition indicating that production of the vowel was higher and more front in this condition. For the lax vowel, F1 was higher and F2 was lower in the congruent condition indicating that pronunciation was lower and further back in the vocal tract. Movement of the tense vowel up and forward and the lax vowel back and down serves to create greater separation of the two vowels within the vowel space. While the native English speakers did not display any priming effects, their F1 and F2 productions also showed a similar pattern wherein the tense vowel was higher and more front and the lax vowel was lower and further back, consistent with Hillenbrand et al. (1995). Thus, the priming effects seen in the Mandarin speakers' productions are consistent with what would be expected if congruent primes had a facilitatory effect. That is, productions following a congruent prime segment were pushed in the direction of more native-like formant frequencies.

Results for back the back vowel contrast showed an overall effect of tensity for F1 when pooled across the two groups wherein the tense vowel had lower F1 measurements than the lax vowel indicating that it was produced higher in the vocal tract. While this is consistent with the patterns for native English speakers found in previous research (Hillenbrand et al., 1995), there was no significant interaction of group or congruency. Indeed, Mandarin speakers produced the tense and lax back vowels with such little distinction (as displayed in Figure 4) that the overall tensity effect seen here was likely driven by the English productions. Additionally, no significant effects were found for F2 in the back vowel contrast.

4.2. Native Listener Judgements

Consistent with the acoustic results which showed Mandarin speakers producing tense vowels higher and more front and lax vowels lower and further back in the congruent condition, intelligibility scores showed a significant priming effect wherein front vowels produced following a congruent prime segment were more accurately identified by native English listeners than those following an incongruent prime. These results seem to suggest that the combined effect of Mandarin speakers producing exaggerated durational differences for the front vowels as well as the more native-like formant frequencies found following congruent primes had a positive effect on the responses of native English listeners, at least in terms of intelligibility. The same cannot be said for the back vowels. However, this is not surprising because both durational cues and formants were likely too dissimilar from native English productions to facilitate more accurate identification by listeners even in congruent conditions. In addition, across vowels and congruency conditions English productions were more accurately identified than Mandarin productions indicating that Mandarin speakers' productions were still not perfectly native-like.

For goodness rating, as expected, utterances produced by native English speakers were more highly rated than those produced by Mandarin speakers regardless of tensity and congruency, and there was a lack of congruency effect for both groups. Thus, while prime segments may have been shown to have an effect on formant frequency in L2 productions and consequently on intelligibility (at least for front vowels) this does not seem to transfer into the realm of acceptability or goodness. Such results seem to reflect previous research by Munro and Derwing (1995) which showed that while native speakers may give accented speech very high ratings on intelligibility (the extent to which an utterance is understood) and even comprehensibility (ease of interpretation), accent rating was highly variable with a large number of utterances being rated as "heavily accented". In fact, even after training, native listeners' ratings of accentedness may not necessarily improve even if other dimensions such as comprehensibility and fluency are positively affected (Derwing, Munro & Wiebe, 1998).

4.3. General Discussion

The aim of this study was to examine whether auditory priming of an English segment spoken by a native English speaker would facilitate more accurate production of that segment for Mandarin learners of English. Results of the effects of priming on duration, formant frequency, and native listener judgements are reviewed in Section 4.3 in light of previous research. Additionally, results were also examined in light of how L2 learners deal with and categorize “familiar” or “similar” versus “novel” or “different” phonetic contrasts as outlined by Flege in the SLM (1995, 2007) and in terms of perceptual assimilation of L2 contrasts into existing L1 categories as outlined by Best and Tyler in the PAM L2 (2007) in Section 4.4.

Results from the front vowel contrast do appear to show a facilitatory effect of the congruent auditory prime for the non-native speakers. While the front vowel contrast was already being readily distinguished by the Mandarin speakers in terms of duration regardless of the presence of an auditory prime as was previously shown to be the case by Wang (1997), formant frequency was positively affected. Mandarin speakers produced more native-like formant frequencies for both tense and lax vowels in the trials containing a congruent prime segment (although F2 for the tense vowel in the congruent condition was not significantly different than that of the incongruent condition, it was heading in a more native-like direction). Additionally, these results are also consistent with previous research indicating that perceptual input from native speakers has a significant facilitatory effect on the productions of L2 learners (Flege, 2003; Flege, 1991; Flege & Eefting, 1987). Accordingly, this shows that exposure to a word spoken by a native speaker preceding pronunciation of that word results in faster response times (Trofimovich & Gatbonton, 2006). Moreover, this suggests that even brief exposure to native speaker input can impact production for nonnative learners, at least on the short-term and for some contrasts.

However, what may initially seem surprising here is that formant frequency in the congruent condition was positively affected for the tense vowel /i/ despite the fact that it exists in the Mandarin vowel inventory and should therefore not have been as readily affected by the prime segment, similar to how native productions were unaffected. One possibility is that the incongruent condition actually caused an inhibitory effect, causing productions of /i/ to be less English-like and even less Mandarin-like. However, if this were the case we would expect English productions and duration to be similarly affected by the incongruent prime. Another explanation is that the Mandarin high front vowel is actually less like the English high front vowel than was expected. If this were the case, we might expect Mandarin speakers to have difficulty producing a native-like high front vowel as predicted in the SLM where “similar” phones prove more difficult for L2 speakers to produce than “new” phones (Flege, 1987). Results from this study indicate that Mandarin productions of /i/ as pooled across congruency conditions are produced lower and further back than the English equivalents and previous research has found average F1 and F2 measurements for Mandarin /i/ that are similar and yet not identical to those for English /i/ found in this study (Li and Xu, 2005; Yang, 1999). Thus, further research may benefit from inclusion of native Mandarin productions of Mandarin vowels as a baseline to compare to Mandarin-produced English vowels.

In terms of native listener judgements, it appears that the Mandarin speakers' predisposition to produce a significant durational difference between the front tense and lax vowels coupled with increased formant accuracy resulted in better pronunciation accuracy, and ultimately intelligibility, in the congruent prime condition as judged by native English listeners. This is consistent with previous clinical research examining cerebral palsy patients, which showed that a positive relationship exists between vowel acoustics and speech intelligibility (Kim, Hasegawa-Johnson & Perlman, 2011; Liu, Tsao & Kuhl, 2005), specifically with regards to formant overlap and vowel space size. Additionally, linguistic research also shows that low intelligibility is correlated with reduced vowel space (Bradlow, Torretta & Pisoni, 1996). Figure 5 above (Section 3.1.2) clearly shows tense and lax high vowels produced by Mandarin speakers in the incongruent condition in much closer proximity in the vowel space than those in the congruent condition indicating that a congruent prime facilitates an expanded vowel space and therefore seems to contribute to increased intelligibility.

On the other hand, back vowels did not seem to be affected by prime segments for any of the four dimensions that were measured. However, as no effect was seen for either of the acoustic measurements then one might reasonably expect that listener judgements would not be affected either and so this at least is not surprising. Still it is not such a simple matter to explain why the acoustic dimension was not affected in the first place when data from the front vowel set clearly showed a facilitatory effect from the congruent prime segments. One reason might be attributable to the minimal pair set used in this study in terms of both orthography and word frequency. At the time of stimuli selection considerations such as consistent orthographical representations of the vowels were weighed against the relative frequency of each of the chosen target words, particularly with regard to the back vowels as there are fewer minimal pair contrasts to choose from for the back vowel contrast. Words that were ultimately chosen as targets appeared no less than 1900 times each according to the Corpus of Contemporary American English. Although not considered to be frequent words, they occurred more frequently than the alternatives (e.g., for the minimal pair “should”/”shooed”, “shooed” appeared only 232 times). Thus future research may benefit from having a larger number and variety of target words to mitigate any potential effects of stimuli selection.

Additionally, goodness rating was also unaffected by congruency despite the fact that prime effects appear to be facilitative for formant frequency and in turn intelligibility for the front vowels. As discussed in Section 4.2 native English listeners rated Mandarin productions equally across congruency conditions and significantly lower than native English productions, a result consistent with previous research that has shown that intelligibility and goodness rating are not necessarily correlated (Munro & Derwing, 1995; Derwing et al., 1998). Priming segments have been shown here to be successful in facilitating more native-like production for Mandarin speakers and thus intelligibility, rather than exacting a global effect on pronunciation.

In terms of native versus non-native results, as discussed above, results revealed that while Mandarin speakers did appear to experience facilitatory effect in producing L2 sounds, native English speakers were not affected by the prime segments for any of the four dimensions that were measured. It could be the case that English participants simply reached ceiling as their perception of the target vowels was near perfect, and they produced acoustically distinctive tense and lax vowels which were judged as perfect as well. This is not surprising since, as native speakers, the English participants have established robust categories for these vowels. In such cases, an auditory prime will be unlikely to affect production. This is consistent with previous priming research that uses the presence of noise (Tilsen, 2009; Nielsen, 2007) or other adverse conditions (Trofimovich, 2005a) to unmask priming effects that might otherwise be hidden by the ceiling effect. In this case, Mandarin speakers' positive responses to the priming tasks for the front vowel are likely a result of the fact that functioning within one's L2 could act as an adverse condition.

In sum, this study revealed that for English speakers prime segments, either congruent or incongruent, were equally ineffective for producing differences in vowel duration, formant frequency, word identification and rating. This was also the case for Mandarin speakers producing the back vowel contrast. However, in the instance of the front vowel contrast, Mandarin speakers were positively affected by congruent prime segments for one of the acoustic measures (formant frequency) and one of the judgement measures (intelligibility) resulting in more native-like pronunciation in such cases. These results are reflected in the body of previous literature looking at the connection between production and perception whereby a clear and measurable connection was established across the two domains (Bradlow et al., 1999; Bradlow et al., 1997; Akahane-Yamada et al., 1996; Rochet, 1995; Pimsleur, 1963; D'Ausilio et al., 2009; Meister et al., 2007; Fadiga et al., 2002; Roy et al., 2008; Cooper & Lauritsen, 1974; Cooper et al., 1975; Kawahara, 1998; Jones & Munhall, 2005; Houde & Jordan, 1998; Purcell & Munhall 2006; Villacorta et al., 2007).

4.4. Theoretical Implications

The current results showing how perceptual information can influence articulation of non-native sounds can be addressed within the framework of the SLM and PAM-L2.

In terms of the perceptual test, results from the Mandarin speakers seem to support the SLM which states that L2 sounds with reduced perceived differences from an L1 sound may be more difficult to discriminate. In this case it does seem that the English tense/lax vowel contrast was somewhat difficult for the Mandarin speakers to discriminate and from a SLM perspective the two separate vowels in the contrast (tense and lax) have likely been perceived as equally similar to the Mandarin tense vowel (Flege, 1995, 2007). A PAM-L2 approach would also be appropriate in explaining the perceptual test results for Mandarin speakers as this theory predicts that two L2 phonological categories (in this case tense and lax vowels) which are perceived as equivalent to the same L1 phonological category will result in discriminatory difficulty and the need for the learner to create a new phonetic category (Best & Tyler, 2007). Results seem to indicate that the Mandarin speakers have perhaps begun to create a new category to encompass the English tense/lax contrast as they did perform better than chance on the perceptual test even if they were not yet at native-like accuracy.

In terms of priming effects, results indicate that perception and production are linked, a result which is predicted by both PAM-L2 and SLM for non-native speakers. From the SLM perspective the production of an L2 sound will be learned as a result of establishing a corresponding perceptual category for that sound. Such a hypothesis requires that pronunciation will always follow perception in terms of accurate acquisition. As mentioned, results from this study do seem to suggest that the Mandarin speakers have at least begun to create a new category for the English tense and lax vowel contrast as they did perform better than chance on the perceptual test. According to the PAM-L2, which is based on direct realism (Fowler 1986), the listener directly perceives articulatory gestures in the speech signal. The findings in this study do seem to support the idea that the perceptual objects of speech are articulatory gestures as pronunciation information in the speech signal did seem to be available to and utilized by the Mandarin listeners to more accurately produce the high front tense and lax vowels.

In sum, results of the current study seem to support the shared view of the SLM and PAM-L2 that perceptually similar L2 contrasts may be assimilated into or indistinguishable from existing L1 phonetic categories making both perception and production a challenge for L2 learners. Additionally, gestural information does appear to be available to L2 learners as a source of information regarding accurate pronunciation of L2 sounds and can also be explained in terms of a gestural account of speech perception as embraced by PAM-L2.

5. Conclusion and Future Directions

The aim of this study was to add to existing research surrounding the connection between the perception and production of speech sounds, particularly for L2 learners and to extend the priming paradigm to examine the cross-modal effects of auditory primes on pronunciation-focused targets. Results showed that although perception for L2 learners on a difficult L2 contrast may not be as accurate as for native speakers, auditory primes can facilitate more native-like pronunciation of that contrast.

The present research also provides useful corroboratory results for accent research, supporting previous studies that revealed that access and exposure to native speakers of one's L2 has a positive impact on pronunciation (Flege, 2003; Flege, 1991; Flege & Eefting, 1987). Furthermore, it supports previous work which has shown that native listeners tend to retain high standards for accentedness or goodness rating even when intelligibility is high (Munro & Derwing, 1995; Derwing et al., 1998). This is useful for L2 learners and teachers when it comes to creating a curriculum for pronunciation teaching, allowing teachers to confidently focus on areas that will exact a higher social benefit (e.g. increasing intelligibility) while not fixating on perfect native-like pronunciation.

The current inconsistent results for the front and back vowels may motivate future research to include additional L2 contrasts to examine how different L2 sounds may be otherwise affected by auditory priming. Results did seem to indicate that the back vowel contrast was more difficult for the Mandarin speakers to produce even across congruency conditions, thus priming effects such as those described in this study may only be observed in L2 contrasts which are already somewhat distinguishable. Furthermore, a separate task wherein an articulatory prime either overt or covert precedes a perceptual target could also shed some light into the connection between perception and production. Thus, this study has opened the door to a number of possible avenues of investigation for examining the link between perception and production as well as having contributed to our understanding of how auditory information is used by L2 speakers to inform and influence their L2 pronunciation.

References

Akahane-Yamada, R., Tohkura, Y., Bradlow, A. R., & Pisoni, D. B. (1996). *Does training in speech perception modify speech production?* Paper presented at *ICSLP 96*. doi: 10.1109/ICSLP.1996.607434

Baese-Berk, M. M. (2010). *An examination of the relationship between speech perception and production.* (Ph.D., Northwestern University). Retrieved from ProQuest Dissertations and Theses. (839790760).

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167-200). Timonium, MD: York.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro, & O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Bohn, O. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279-304). Baltimore: York Press.

Bradlow, A., Pisoni, D., Akahane-Yamada, 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*, 2299-2310.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61*(5), 977-985.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20*(3-4), 255-272.

Chen, Y. (1999). *Acoustic characteristics of American English produced by native speakers of Mandarin.* (Ph.D., University of Connecticut. Retrieved from Dissertation Abstracts International, B: Sciences and Engineering. (200010573).

Cooper, W. E., & Lauritsen, M. R. (1974). Feature processing in the perception and production of speech. *Nature, 252*, 121-123.

Cooper, W. E., Blumstein, S. E., & Nigro, G. (1975). Articulatory effects on speech perception: A preliminary report. *Journal of Phonetics, 3*(2), 87-98.

D'Ausilio, A., Pulvermuller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology, 19*(5), 381-385.

Davies, M. (2008-). *The corpus of contemporary American English: 450 million words, 1990-present.* Retrieved from http://corpus.byu.edu/coca/.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393-410.

Dreher, B., & Larkins, J. (1972). Non-semantic auditory discrimination: Foundation for second language learning. *The Modern Language Journal, 56*(4), 227-230.

Elman, J. L. (1981). Effects of frequency- shifted feedback on the pitch of vocal productions. *The Journal of the Acoustical Society of America, 70*(1), 45-50.

Fadiga, L., Craighero, L., Buccino, G., & Rizzollati, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience, 15*(2), 399-402.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole, & J. I. Hualde (Eds.), *Laboratory Phonology 9* (pp. 353-382). Berlin: Mouton de Gruyter.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233-277). Timonium, MD: York Press.

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

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

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. Schniller, & A. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319-355). New York: Mouton de Gruyter.

Flege, J. E., Bohn, O., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics, 8*(2), 162-177.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America, 97*(5), 3125-3134.

Fowler, C., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language, 49*(3), 396-413.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics, 14*(1), 3-28.

Garcia Perez, G. M. (2003). *Training Spanish speakers in the perception and production of English vowels.* (Ph.D., Simon Fraser University). Retrieved from ProQuest Dissertations and Theses. (NR03118).

Hattori, K. (2010). *Perception and production of English /r/-/l/ by adult Japanese speakers.* (Unpublished doctoral dissertation). University College London, London.

Hernandez, A. E., Fennema-Notestine, C., Udell, C., & Bates, E. (2001). Lexical and sentential priming in competition: Implications for two-stage theories of lexical access. *Applied Psycholinguistics, 22*(2), 191-215.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*(5), 3099-3111.

Hirata, Y., & Whiton, J. (2005). Effects of speaking rate on the Single/Geminate stop distinction in Japanese. *The Journal of the Acoustical Society of America, 118*(3), 1647-1660.

Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes, 8*(4), 379-411.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science, 279*(5354), 1213-1216.

Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory & Cognition, 12*(6), 568-574.

Jones, J. A., & Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Current Biology, 15*(19), 1768-1772.

Kawahara, H. (1998). Hearing voice: Transformed auditory feedback effects on voice pitch control. *Proceedings of the International joint conference on artificial intelligence: Workshop on computational auditory scene analysis* (335-349). Mahwah, NJ: Erlbaum.

Kim, H., Hasegawa-Johnson, M., & Perlman, A. (2011). Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatrica Et Logopaedica, 63*(4), 187-194.

Kondaurova, M., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America, 124*(6), 3959-3971.

Kozhevnikov, V. A., & Chistovich, L. A. (1965). *Speech: Articulation and perception* (revised ed.). Washington: U.S. Department of Commerce.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather, & A. James (Eds.), *New sounds 90* (pp. 72-97). Amsterdam: University of Amsterdam.

Li, X., & Xu, B. (2005). Formant comparison between whispered and voiced vowels in Mandarin. *ACTA Acustica United with Acustica, 91*(6), 1079-1085.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

Liu, H., Tsao, F., & Kuhl, P. K. (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *The Journal of the Acoustical Society of America, 117*(6), 3879-3889.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*(3), 251-285.

Meister, I., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology, 17*, 1692-1696.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisiton, 21*(1), 81-108.

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

Neufeld, G. G. (1980). On the adult's ability to acquire phonology. *TESOL Quarterly, 14*(3), 285-298.

Neufeld, G. G. (1988). Phonological asymmetry in second-language learning and performance. *Language Learning, 38*(4), 531-559.

Nielsen, K. Y. (2007). The interaction between spontaneous imitation and linguistic knowledge. *University of California Working Papers in Phonetics, 105*(Mar), 125-137.

Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I. P., Joensuu, R., Autti, T., & Sams, M. (2005). Processing of audiovisual speech in Broca's area. *Neuroimage, 25*(2), 333-338.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*(3), 261-283.

Peperkamp, S., & Bouchon, C. (2011). The relation between perception and production in L2 phonological processing. Paper presented at the *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011),* 161-164.

Perez-Gamboa, T. (2001). *An investigation of the production and perception of an English phonological contrast /i/ and /I/: The effects of pronunciation training on the speaking and listening skills of Peruvian learners of English as a foreign language.* (Ph.D., State University of New York at Stony Brook). Retrieved from ProQuest Dissertations and Theses. (9989388).

Pilotti, M., Bergman, E. T., Gallo, D. A., Sommers, M., & Roediger, H. L.,III. (2000). Direct comparison of auditory implicit memory tests. *Psychonomic Bulletin & Review, 7*(2), 347-353.

Pimsleur, P. (1963). Discrimination training in the teaching of French pronunciation. *The Modern Language Journal, 47*(5), 199-203.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*(2), 191-215.

Porter, R. J., & Castellanos, F. X. (1980). Speech-production measures of speech perception: Rapid shadowing of VCV syllables. *The Journal of the Acoustical Society of America, 67*, 1349-1356.

Porter, R. J., & Lubker, J. F. (1980). Rapid reproduction of vowel-vowel sequences: Evidence for a fast and direct acoustic-motoric linkage in speech. *Journal of Speech and Hearing Research, 23*, 593-602.

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America, 120*, 966.

Random.org. (2012). Retrieved July/7, 2013, from http://www.random.org/lists

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 379-410). Timonium, MD: York Press.

Rogers, H. (2000). *The sounds of language: An introduction to phonetics*. Harlow: Longman Publishing Group.

Roy, A. C., Craighero, L., Fabbri-Destro, M., & Fadiga, L. (2008). Phonological and lexical motor facilitation during speech listening: A transcranial magnetic stimulation study. *Journal of Physiology-Paris, 102*(1–3), 101-105.

Schneiderman, E., Bourdages, J., & Champagne, C. (1988). Second-language accent: The relationship between discrimination and perception in acquisition. *Language Learning, 38*, 1-19.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America, 125*, 1103.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *Neuroimage, 25*(1), 76-89.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex, 17*(10), 2387-2399.

Spencer, K. A., & Wiley, E. (2008). Response priming patterns differ with interstimulus interval duration. *Clinical Linguistics & Phonetics, 22*(6), 475.

Tilsen, S. (2009). Subphonemic and cross-phonemic priming in vowel shadowing: Evidence for the involvement of exemplars in production. *Journal of Phonetics, 37*(3), 276-296.

Trofimovich, P. (2005a). Spoken-word processing in native and second languages: An investigation of auditory word priming. *Applied Psycholinguistics, 26*(4), 479-504.

Trofimovich, P. (2005b). Implicit memory support for bilingual speech processing: An investigation of auditory word priming in English-Spanish bilinguals. *Proceedings of the 4th International Symposium on Bilingualism.* 2257-2274.

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal, 90*(4), 519-535.

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America, 122*, 2306.

Wang, X. (1997). *The acquisition of English vowels by Mandarin ESL learners: A study of production and perception.* (Unpublished Doctoral dissertation). Simon Fraser University, Vancouver.

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

Wang, X. (2002). *Training Mandarin and Cantonese speakers to identify English vowel contrasts: Long-term retention and effects on production.* (Ph.D., Simon Fraser University). Retrieved from ProQuest Dissertations and Theses. (NQ81699).

Watkins, K., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience, 16*(6), 978-987.

Werker, J. F. (1989). Becoming a native listener. *American Scientist, 77*(1), 54-59.

Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*(1), 349-355.

Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage, 33*(1), 316-325.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience, 7*(7), 701-702.

Appendices

Appendix A.

Target and distractor sentences for the inclusion test

Sentences including target words	Sentences including distractor words
The farmer raises sheep.	We should head back.
He is the captain of the ship.	Janet says she can't come.
The boy performed a good deed.	Melt four tablespoons of bacon fat.
I didn't steal it, he *did*.	Greg was trying to make a quick buck.
Horses have four feet.	My brothers always fought.
The shoe does not fit.	Try to remain calm.
Sally wants to eat a peach.	
At baseball camp, I learned to pitch.	
You're acting like a drunken fool.	
The bucket is full.	
My father's name is Luke.	
During the scary parts, I don't look.	
Let's visit the pool.	
A good dog doesn't pull.	
Who'd like some ice cream?	
John has a jacket with a hood.	