Spring 2013

Perception and production of English cues to plosive voicing by native Mandarin speakers

Jessica Lutz
University of Colorado Boulder

Perception and production of English cues to plosive voicing by native Mandarin speakers

Jessica Lutz

Linguistics Departmental Honors Thesis
University of Colorado at Boulder

April 8, 2013

Thesis Advisor
Rebecca Scarborough | Assistant Professor | Linguistics

Committee Members
Eliana Colunga | Assistant Professor | Psychology & Neuroscience
Kira Hall | Associate Professor | Linguistics & Anthropology

Abstract

The acquisition of non-native sounds is a prominent feature of second language acquisition. Mandarin does not permit the voiced plosives /b d g/ or plosives in the syllable coda position, both features of many English words. This disparity presents a potential barrier to native Mandarin speakers' perception and production of English. Native English speakers use two cues to determine word-final plosive voicing: vowel duration differences and voicing of the closure itself. To assess native Mandarin speakers' acquisition of English word-final voiced plosives, their perception and production of English cues to plosive voicing were examined through acoustical analyses and a forced-choice identification task. Native Mandarin speakers were not able to reliably produce voiced closures, but did produce significant vowel duration differences. In addition, native Mandarin speakers identified voiced plosives consistently with the voicing of the closure, while native English speakers identified voiced plosives consistently with vowel duration. These findings can be explained with the hypotheses of language transfer and the Speech Learning Model.

Table of Contents

1 Introduction

The acquisition of a second language can be an arduous process. Second language learners often produce phonemes unlike the norms expected by native speakers, resulting in accented speech. The effects of accents range from hindered communication to the development of negative stereotypes, but they also provide insight into the acquisition of non-native sounds.

In this study, the "Chinese accent" is examined by focusing on native Mandarin speakers' production and perception of English cues to plosive voicing. This introduction surveys applicable hypotheses of second language acquisition and provides a description of the relevant phonological characteristics of Mandarin and English.

1.1 Relevant Hypotheses of Second Language Acquisition

The field of second language acquisition (SLA) involves the acquisition of all elements of a second language. Most of the following hypotheses can be applied to multiple levels of acquisition; however, for the purposes of this study, they will be analyzed with respect to the acquisition of non-native phonology.

1.1.1 Language Transfer

Language transfer is one of the earliest theories of SLA. In the study of phonological acquisition, transfer is defined as the tendency to substitute a novel second language (L2) sound with a similar first language (L1) sound (Archibald, 1998). The effect of this transfer is divided into two categories. Positive transfer occurs when the L1 and L2 patterns are similar, and habits from the L1 are successfully carried over to the L2. Negative transfer occurs when the L1 and L2 patterns are different, and habits carried over interfere with the L2 (Ellis, 1994; Major, 2008).

The Contrastive Analysis Hypothesis (CAH) arose from the idea of language transfer. Popular in the 1960s, the main tenet of the CAH is that greater differences between the L1 and L2 increase the difficulty of L2 acquisition. Subsequent research demonstrated that differences between the L1 and L2

are not always directly predictive of errors, because speakers often go through similar developmental stages in L2 acquisition regardless of the L1 (Ellis, 1994; Loewen, 2011). The CAH is no longer considered a prominent theory of SLA, but the role of transfer and L1 influence on L2 are still widely accepted.

1.1.2 Markedness

The concept of markedness assumes that there are language universals, or trends that hold true across all languages. Markedness exists on multiple feature levels, from phonemes to syntactical structures. "Unmarked" forms are common, prototypical forms found in many languages. "Marked" forms are rare and often complex. For instance, the sound [m] is considered unmarked, because it is a common sound found in many languages, while the less common [ð] is considered marked (Hansen, 2006; Ortega, 2009).

The Markedness Differential Hypothesis (MDH), developed by Fred Eckman in 1977, posits that marked forms are more difficult to acquire than unmarked forms. Additionally, it suggests that marked L1 forms are less likely to be transferred than unmarked L1 forms (Eckman, 1977).

1.1.3 The Speech Learning Model

The Speech Learning Model (SLM) is a prominent modern theory of SLA proposed by James Flege. Although individual hypotheses of the SLM have changed dramatically since it was popularized in the 1990s, it endures as one of the principal models relating the perception and production of non-native sounds. One distinguishing characteristic of the SLM is that it assumes transfer occurs bidirectionally; that is, the L2 influences the L1 just as the L1 influences the L2. The SLM also rejects the common belief that there is a "critical period" to language learning, an age range after which it is almost impossible to acquire native phonology (the Critical Period Hypothesis). Instead, the SLM suggests that adults are able to continue learning non-native phonology throughout life (Flege, 1995; Flege 2003).

The SLM includes many hypotheses, and two are relevant to the present study. The first relevant hypothesis presents the idea of category formation: that sounds in the L2 are grouped into categories

based on perceived similarities to L1 sounds. Less similar sounds are more likely to form new categories. The second hypothesis, which will be called "the assimilation hypothesis of the SLM" for the purposes of this study, states that L2 sounds perceived similarly to L1 sounds will not form a new category. The L1 and L2 will instead assimilate, creating a merged L1-L2 category. Bi-directional transfer will occur within the category, changing both the L1 and L2 sounds (Flege, 2003).

1.2 Native Mandarin Speakers' Acquisition of English

The phonology of Mandarin Chinese differs dramatically from that of English, from the basic consonant inventory to underlying phonological processes. This study addresses differences between English and Mandarin plosive inventories and syllable structures, as well as English cues to plosive voicing.

1.2.1 Relevant Mandarin Phonology

The Mandarin consonant inventory contains the voiceless plosives /p^h p t^h t k^h k/ in contrastive distribution. Mandarin contains no voiced plosives. Although Mandarin allows both CVC and CV syllable structures, CV is most common, and CVC allows only nasal codas. Mandarin plosives are therefore exclusively realized in the syllable onset position (Duanmu, 2007; Li, 1999; Xu, 2001). Mandarin does not contain contrastive vowel duration (Shen, 1991).

1.2.2 Relevant English Phonology

The English consonant inventory includes the voiced and voiceless plosives /p b t d k g/ in contrastive distribution. English also permits the aspirated voiceless plosives /p^h t^h k^h/, but these are allophones of the voiceless plosives /p t k/. The CVC syllable structure is very common in English, and voiced and voiceless plosives are often realized as codas (Hammond, 1999).

English also contains the feature of vowel duration as a cue to plosive coda voicing. Vowels appearing before voiced plosive codas have characteristically longer durations than vowels before voiceless plosive codas. A study performed by Lawrence Raphael in 1971, as well as many studies since,

demonstrated that native English speakers discriminate voiced and voiceless plosives consistently with vowel duration regardless of the voicing of the closure itself (Raphael, 1971).

1.2.3 Potential Barriers to English Acquisition

From the basic perspective of language transfer and L1 influence on L2, Mandarin's lack of voiced plosives /b d g/ and plosive codas is likely to hinder native Mandarin speakers' acquisition of the English voiced and voiceless plosive codas. This difficulty may be compounded with respect to the MDH, because both voiced plosives and plosive codas are considered marked, and therefore difficult to acquire.

The vowel duration cue to plosive voicing may be a less difficult acquisition. Previous studies have shown that vowel duration is learned more quickly than coda voicing (Flege, 1992). In addition, children produce the vowel duration cue to closure voicing before they are able to produce voicing in the closure itself (Naeser, 1970).

Most research into native Mandarin speakers' acquisition of English examines plosives in the syllable onset position, but some is known about acquisition of word-final plosives. Perhaps the most formative study was published in 1992 by James Flege, assessing the production of word-final English /t/ and /d/ by native English, Mandarin, and Spanish speakers. He found that native Mandarin speakers were able to produce small closure voicing contrasts, as well as significant vowel duration contrasts. Both contrasts were less dramatic than those of native English speakers, and perception was not addressed (Flege, 1992).

1.3 Current Study

This study assesses native Mandarin speakers' production and perception of the non-native cues to word-final plosive voicing: vowel duration and closure voicing itself.

The study is divided into two parts: a production study and a perception study. The production study serves to confirm that Flege's findings regarding native Mandarin speakers' production of /t/ and /d/ apply to the subjects in this study, as well as extend to /p b k g/. Native Mandarin speakers' production of

both English cues to plosive voicing is examined with acoustic analyses. The perception study then assesses native Mandarin speakers' interpretation of contradictory cues to plosive voicing through a forced-choice identification task. Conclusions are made about the relationship between perception and production, as well as how these findings fit into relevant theories of SLA.

2 Production Study

Native Mandarin speakers' accuracy in production of both standard English cues to plosive voicing were assessed through acoustic analysis of minimally paired CVC words containing the English word-final plosives /p b t d k g/. Measurements of closure voicing and vowel durations were compared to those of native English speakers.

2.1 Methods

2.1.1 Subjects

Fourteen subjects living in Boulder, CO were recruited by personal contact. The seven native Mandarin speakers, four male and three female, had a mean age of 22 years. All had learned English as an L2 after the age of 14, and moved to the United States from mainland China after the age of 18. Two native Mandarin speakers were bilingual speakers of Cantonese, which shares phonological characteristics with Mandarin relevant to this study. The seven native English speakers, three male and four female, had a mean age of 21 years. All were monolingual English speakers from the western United States.

2.1.2 Procedure

Speakers were recorded while reading a list of forty CVC words. All words contained one of the vowels /ɪ æ ɑ ʌ/. The list was randomized and included eight /CVp/ /CVb/ minimal pairs, eight /CVt/ /CVd/ minimal pairs, and eight /CVk/ /CVg/ minimal pairs. Words were assessed for familiarity using the Hoosier mental lexicon; only words with familiarity scores of five or above were used (Nusbaum et al., 1984). For a complete list of the minimally paired words, see Appendix A.

Both acoustic analyses were performed using Praat software. Closure voicing duration was measured from the point of constriction to the final peak of sinusoidal voicing in the stop closure. Constriction was defined as a sharp decrease in amplitude; meaning vocal tract closure had removed

frequencies above the F1 region (Flege et al., 1992). Measurements of closure voicing duration were made in words with voiced closures, for a total of 280 measurements. Vowel duration was measured from the onset of the vowel's first formant to the constriction of the word-final plosive. Measurements of vowel duration were made in all forty words from each subject, for a total of 560 measurements.

2.2 Results

Two cues to plosive voicing were analyzed: closure voicing duration and vowel duration. Closure voicing duration was examined to determine the ability of native Mandarin speakers to produce non-native sounds appearing in non-native contexts. Vowel duration was then examined with respect to its environment; specifically, differences between vowels preceding word-final voiced or voiceless plosives.

2.2.1 Closure Voicing Duration

Voicing duration of voiced plosive closures was considerably longer when produced by native English speakers than by native Mandarin speakers, averaging 76 ms and 7 ms, respectively. These averages are given in Figure 2.1. A within-subject t-test revealed this difference was significant, ($p<.001$). Mean voiced closure durations produced by each speaker were also calculated. Notably, all seven of the native Mandarin speakers produced closure voicing in under half of the "voiced closure" words, and produced no closure voicing at all in the remaining words. Native Mandarin speakers therefore appeared to have substantial difficulty producing voiced closures when compared to native English speakers.

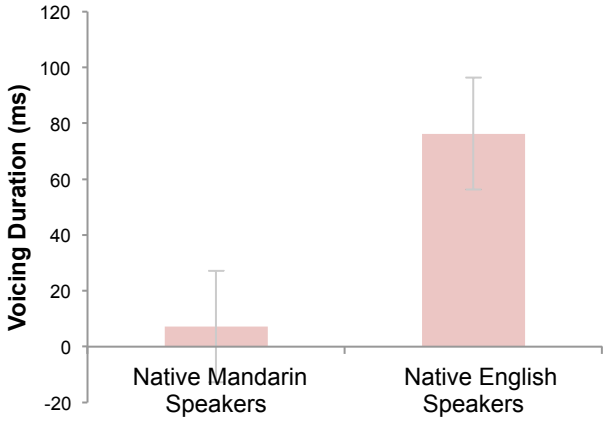

Figure 2.1. Mean voicing duration of voiced plosive closures.

2.2.2 Vowel Duration

Both native Mandarin speakers and native English speakers produced longer vowels in words with voiced plosive closures than voiceless plosive closures. The difference in mean vowel duration produced by native Mandarin speakers in these two contexts averaged 84 ms (32%). Native English speakers produced a greater difference in mean vowel duration, averaging 137 ms (42%). Interestingly, the difference between speaker groups was predominantly due to the durations before voiced closures, as both native English and native Mandarin speakers produced vowels of similar mean duration before voiceless closures. The mean durations of vowels produced by speaker groups in both closure contexts is shown in Figure 2.2.

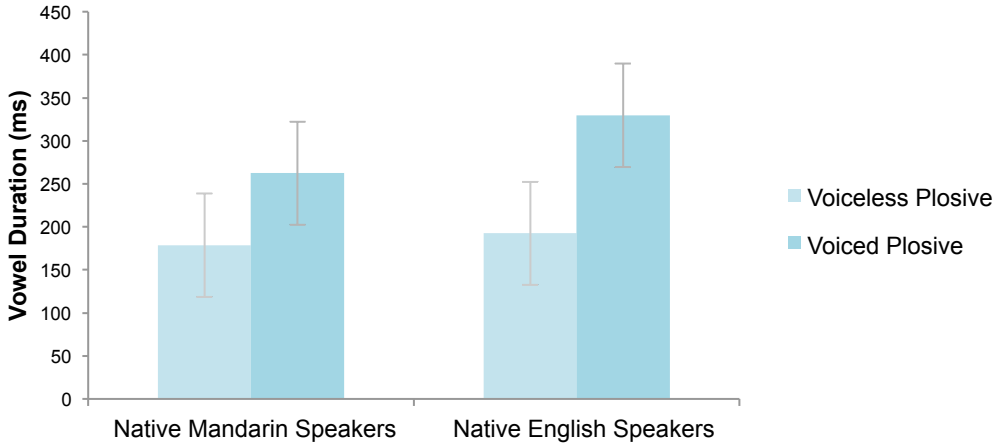

Figure 2.2. Mean duration of vowels produced by native English and native Mandarin speakers before voiced and voiceless plosives. Both groups produced longer vowels before voiced plosives.

A two-way ANOVA (language group x plosive context) was performed on the durations of vowels. This revealed a significant main effect of language group, meaning native English speakers produced longer vowels than native Mandarin speakers ($F(1,554)=51.706$, $p<.001$). A significant main effect of plosive context was also found, indicating vowels before voiced plosives were significantly longer than before voiceless plosives ($F(1,554)=383.160$, $p<.001$). The analysis also revealed a significant language group x plosive context interaction ($F(1,554)=22.528$, $p<.001$). The difference in

vowel duration produced by native English speakers was therefore significantly greater than the difference produced by native Mandarin speakers.

Within-subject t-tests served as post-hoc analyses. They revealed the duration differences within each speaker group were significant, ($p<.001$ for both groups). Native Mandarin speakers were clearly able to produce contrasting vowel lengths related to closure voicing, but not as efficiently as native English speakers.

2.3 Discussion

The production study yielded two basic conclusions. First, native Mandarin speakers were unable to reliably produce voiced plosive closures matching those of native English speakers. Second, native Mandarin speakers were able to produce contrasting vowel durations related to plosive voicing, although not as effectively as native English speakers. These findings are consistent with the Flege's previous findings stated in the introduction (Flege, 1992).

One potential reason for native Mandarin speakers' inability to reliably produce voiced plosive closures is language transfer. Recall, Mandarin lacks voiced plosives, but does contain the voiceless plosives /p^h p t^h t k^h k/ in contrastive distribution. Native Mandarin speakers may associate these phonemes with the English /p b t d k g/, respectively. If this association results in substitution of native plosives for distinct non-native plosives, negative transfer occurs. The plosives produced would be distinct, but not with respect to voicing. For example, the English /p/ could be realized as the Mandarin /p^h/, and the English /b/ realized as the Mandarin /p/. This hypothesis could be tested by comparing native Mandarin speakers' production of these phonemes in Mandarin and in English.

Another potential reason for the lack of plosive closure voicing is provided by the assimilation hypothesis of the Speech Learning Model. This hypothesis is similar to that of language transfer, but also assumes L2 influence on L1. If native Mandarin speakers did not form new categories for the English voiced plosives due to their similarities to the Mandarin plosives /p t k/, both categories would assimilate under the SLM. Therefore, native Mandarin speakers would also produce a non-native sound in

Mandarin, a merged L1-L2 sound. This hypothesis is much more difficult to assess, because the plosives produced by native Mandarin speakers before learning English would need to be included in the above comparison assessing language transfer.

The results of the production study can also be interpreted with to the Markedness Differential Hypothesis. First, native Mandarin speakers were able to produce the vowel duration cue to closure voicing much more effectively than closure voicing itself. This finding supports the MDH because voiced plosives are marked sounds, and are therefore considered difficult to acquire. Additionally, plosives in syllable-final position are marked, making a non-native syllable-final voiced plosive a doubly difficult sound to produce. Vowel duration, on the other hand, is merely a characteristic of sounds native Mandarin speakers can already produce. As mentioned in the introduction, previous studies also support the acquisition of vowel duration cues before closure voicing itself (Raphael, 1971).

It is important to note that, although the native Mandarin speakers demonstrated contrasting vowel durations by lengthening vowels "before voiced plosives," speakers were not actually producing the voiced plosive after the vowel. This indicates that the lengthened vowel was not a result of the voicing of the following plosive. Instead, speakers were ostensibly able to detect vowel duration as a cue to voicing and produce the contrast. This possibility was further explored during the perception study.

3 Perception Study

Native Mandarin speakers' interpretation of contradictory cues to word-final plosive voicing was assessed using a forced-choice identification task. Subjects were presented with modified and unmodified trials, containing words with contradictory and consistent cues, respectively. Subjects' response times to different trials were measured to determine their potential awareness of the contradictory cues, and trends in cues used for word identification were evaluated to determine which cues were significant to native English and native Mandarin speakers.

3.1 Methods

3.1.1 Procedure

To create the words used in the forced-choice identification task, a native English speaker (not a subject) was recorded while reading a list of the same forty CVC words used in the production study. Again, all words contained one of the vowels /ɪ æ ɑ ʌ/, and included eight /CVp/ /CVb/ minimal pairs, eight /CVt/ /CVd/ minimal pairs, and eight /CVk/ /CVg/ minimal pairs. These forty words were used in the "unmodified" trials. The words were then manipulated using Praat software to contain contradicting cues to closure voicing for the "modified" trials. Recall, the two cues to plosive voicing examined in this study were vowel duration and the voicing of the closure itself, and an English word containing a word-final voiced plosive typically contains a lengthened vowel and a voiced closure. Words were therefore manipulated to contain either a lengthened vowel and a voiceless closure or a short vowel and a voiced closure. For example, unmodified trials for the minimal pair /kæp/ and /kæb/ contained [kæp] or [kæ:b], and modified trials contained [kæ:p] or [kæb]. For a complete list of modified and unmodified words, see Appendix B.

The same subjects from the previous study participated in the forced-choice identification task. Psyscope software was used to create the task. Subjects were presented with all eighty randomized modified and unmodified words. As each word was presented, two possible identifications were shown

on the computer screen. For instance, if a subject heard any of the modified or unmodified versions of /kæp/ and /kæb/ listed above, their screen would show the options "cap" and "cab". Subjects were asked to identify which word they heard by pressing a corresponding computer key, and to guess if unsure. Intervals between trials were 700 ms.

Response time measurements were made from the end of the word to the computer key press. RTs above or below two standard deviations from the mean of each speaker group were removed before data analysis to control for key press errors. Key presses were recorded and assessed for trends in cues to word identification.

3.2 Results

Analysis of the forced-choice identification task was divided into two parts. The first part involved subjects' response times. The difference between subject groups' response times in modified and unmodified trials was examined. Observed trends between native English speakers and native Mandarin speakers were assessed for significance.

The second part of the analysis involved only modified trials. In these trials, subjects heard words with contradictory cues to closure voicing, meaning lengthened vowels before voiceless plosives and shortened vowels before voiced plosives. Trends between the identifications of native English speakers and native Mandarin speakers were examined.

3.2.1 Response Times in the Forced-Choice Identification Task

Both native Mandarin speakers and native English speakers appeared to respond more quickly in unmodified trials than in modified trials, shown in Figure 3.1. A two-way ANOVA (language group x modified/unmodified trial) was performed on the RTs, and showed significant main effects for both language group ($F(1,973)=77.691$, $p<.001$) and modified/unmodified trial ($F(1,973)=21.858$, $p<.001$). These imply that English speakers responded more quickly overall, and that modified words had higher RTs across both language groups, respectively. The analysis also showed the language group x

modified/unmodified trial interaction was not significant ($F(1,973)=3.647$, $p>.05$). Taken with the significant main effects, this demonstrates that the native Mandarin and native English speakers' RTs exhibited the same trends. Assuming that the heightened RT in modified trials was a result of the presence of contradictory cues to closure voicing, this analysis demonstrated that, like native English speakers, native Mandarin speakers were able to perceive the contradictory cues in the modified trials.

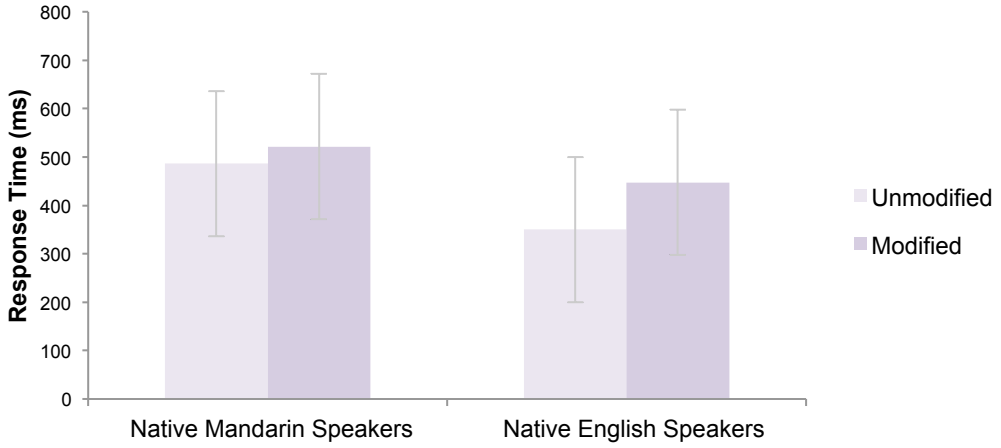

Figure 3.1. Mean response times of native Mandarin speakers and native English speakers during unmodified and modified trials.

3.2.2 Word Identification Cues: Vowel Duration and Closure Voicing

Native English speakers responded consistently with the vowel duration cue, and not with the voicing of the closure itself, in 83% of the modified forced-choice identification trials. This data is consistent with previous findings in native English speaker studies (Raphael, 1971). In contrast, native Mandarin speakers made judgments consistent with vowel duration in only 30% of the modified trials, responding consistently with closure voicing in the remaining 70%. The mean numbers of identifications made by speaker groups based on each cue are shown in Figure 3.2.

A two-way ANOVA (language group x word identification cue) was performed on the number of identifications made in modified trials. The analysis yielded no significant main effect of language group ($F(1,24)=0$, $p=1$), but did yield a significant main effect of word identification cue ($F(1,24)=34.256$, $p<.001$). This finding indicates that, overall, more responses were consistent with vowel duration than

closure voicing. The analysis also revealed a significant language group x word identification cue interaction ($F(1,24)=563.419$, $p<.001$), meaning that there was a significant difference between the cues with which native Mandarin and English speakers responded consistently.

Post-hoc t-tests were performed to confirm that the differences in response trends within each speaker group were significant ($p<.001$ for both). This analysis shows Native Mandarin speakers responded less consistently with vowel duration than closure voicing, while native English speakers demonstrated the opposite trend.

A notable control in this study was the accuracy of responses to the unmodified words. On average, native English speakers misidentified only one of the forty unmodified words, and native Mandarin speakers misidentified four. This proves that both groups were able to use closure-voicing cues effectively in "normal" contexts.

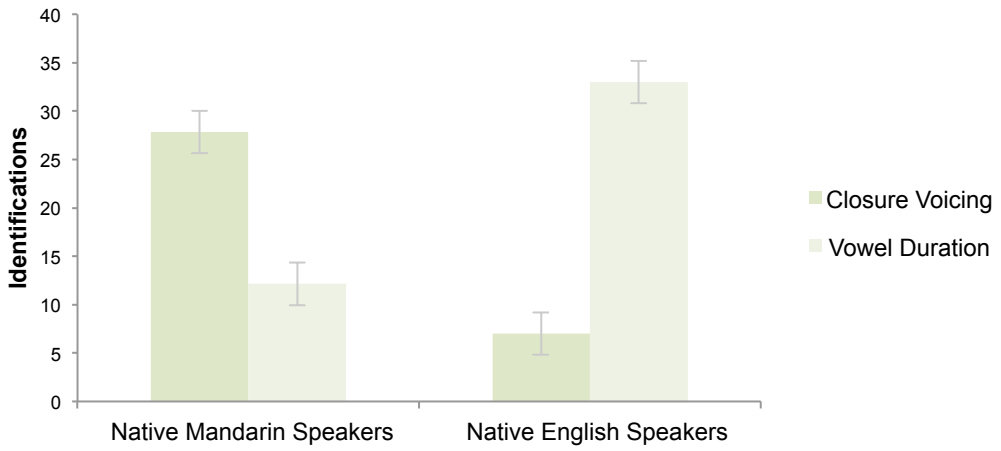

Figure 3.2. Mean number of identifications made by native Mandarin speakers and native English speakers based on closure voicing or vowel duration.

3.3 Discussion

The modified and unmodified trial analysis demonstrates that native Mandarin speakers respond to both cues to plosive voicing in English. This conclusion was drawn under the assumption that native English speakers took longer to respond during the modified trials because they perceived contradicting cues. The fact that there was no significant interaction between speaker group and modified/unmodified

trials indicates native Mandarin speakers also perceived these cues as contradictory, and are therefore aware (on some level) of both cues to closure voicing.

The second analysis revealed a significant difference between the perception of plosive voicing by native Mandarin and English speakers. When presented with contradictory cues, native English speakers appeared to rely on vowel duration to identify words with voiced plosives, while native Mandarin speakers appeared to rely on the voicing of the closure itself. Combined, the analyses show that the native Mandarin speakers were aware of the contradictory cues and were affected by vowel duration, but ultimately chose closure voicing.

The ability of native Mandarin speakers to identify plosives based on voiced closure cues is potentially explained by transfer and the SLM. As mentioned in the production study discussion, negative transfer or assimilation could potentially occur between the Mandarin plosives /p^h p t^h t k^h k/ and the English plosives /p b t d k g/. If native Mandarin speakers consistently perceive these plosives accordingly, there will be no barrier to "accurate" identifications.

The MDH does not provide a reasonable explanation for these results, because native Mandarin speakers identified words based on the more difficult cue. This does not, however, invalidate the MDH. If the results are explained by transfer or the SLM, native Mandarin speakers are not even addressing the issue of voiced plosives. Instead, they are producing and perceiving six voiceless phonemes. The MDH would therefore not apply to these findings.

If the MDH is irrelevant to this study, closure voicing is no longer considered a more difficult cue. But why was it consistently chosen over vowel duration by native Mandarin speakers, while native English speakers made the opposite choice? As discussed in the introduction, cues within plosives are often absent in English, prompting native English speakers to rely on the more consistent cue of vowel duration (Raphael, 1971). These processes are irrelevant in Mandarin, as Mandarin lacks plosives in word-final positions, so native Mandarin speakers rely on the more obvious cue to identify the final segment, the segment itself.

These conclusions are very interesting with respect to those of the production study. Comparisons between the studies, as well as their potential implications for the interaction between perception and production, will be explored in the general discussion.

4 General Discussion

4.1 Perception and Production

At first glance, the results of the production and perception studies may appear to be contradictory. Native Mandarin speakers produced vowel duration differences more effectively than closure voicing, but identified plosives based on closure voicing rather than vowel duration. If these findings are interpreted using the MDH, we must assume that markedness presents a different barrier to production than perception. Essentially, we'd assume that voiced plosives are more easily perceived than produced, or that perception and production are not related.

A much more plausible explanation involves language transfer and/or the assimilation hypothesis of the SLM. (Further analyses are required to distinguish between the two hypotheses with respect to this study, see "Future Directions".) As described in the perception study discussion, if we use these models, the MDH is not invalidated, just irrelevant to the study. Taking into account the influence of perception on production posited by the SLM, we can interpret the study results with the following hypotheses:

1. Native Mandarin speakers perceive the English plosives /p b t d k g/ as /p^h p t^h t k^h k/ through language transfer, or assimilate the two categories as predicted by the SLM.
2. As a result, native Mandarin speakers produce /p b t d k g/ as /p^h p t^h t k^h k/ through language transfer, or produce merged sounds as predicted by the SLM.
3. Native Mandarin speakers can both perceive and produce changes in vowel duration associated with closure voicing.
4. Native English speakers rely on vowel duration cues to closure voicing because cues within closures are often unreliable; Native Mandarin speakers rely on the more obvious cue of the voicing of the closure itself because they are unaccustomed to the unreliability of closure cues.

4.2 Future Directions

In order to confirm hypotheses 1 and 2, similarities between the L1 and L2 plosives must be observed. Initially this will require acoustic analysis of the Mandarin plosives /p^h p t^h t k^h k/. Perceived similarities between closure duration, release, and aspiration may then indicate the process of language transfer or category assimilation. Discriminating between the two will be more challenging, because an analysis of L1 plosives before the native Mandarin speakers acquired English is necessary. If the L1 sound also exhibits a change, category assimilation could be responsible.

Another interesting direction is an examination of the effects of syllable position on acquisition. This would be most easily assessed with voiced and voiceless plosives in syllable onsets. Mandarin allows voiceless plosives in this position, so transfer could again be considered. Additionally, the MDH could be revisited, because it makes predictions regarding the markedness of plosives in onsets and codas.

In conclusion, both the perception and the production studies provide a good foundation for the application of language transfer and the SLM to native Mandarin speakers' acquisition of English. Confirmation of the above hypotheses could provide solid support for these models, as well as further insight into the relationship between production and perception.

Acknowledgements

Above all, I would like to thank Rebecca Scarborough for providing me with invaluable insight into research in phonetics and phonology. This project would not exist without her time, patience, and friendly encouragement. I would also like to thank the study participants, busy students who kindly spent time assisting me with nothing in return. Thanks also to my thesis committee, Eliana Colunga and Kira Hall, for taking time to advise me on this project. Finally, I would like to thank my friend, Yi Bu, who helped inspire this thesis. Her advice concerning Mandarin was essential to these studies and many more during my undergraduate career.

References

Archibald, J. *Second Language Phonology*. Philadelphia, PA: John Benjamins Publishing Company, 1998.

Duanmu, S. *The Phonology of Standard Chinese*. 2nd ed. New York: Oxford University Press, 2007.

Eckman, F. "Markedness and the Contrastive Analysis Hypothesis." *Language Learning* 27 (1977): 315-30.

Eckman, F. "Typological Markedness and Second Language Phonology." In *Phonology and Second Language Acquisition*, edited by Jette G. Hansen Edwards and Mary L. Zampini. Philadelphia, PA: John Benjamins Publishing Company, 2008.

Ellis, R. *The Study of Second Language Acquisition*. New York, NY: Oxford University Press, 1994.

Flege, J. "Assessing Constraints on Second-Language Segmental Production and Perception." In *Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities*, edited by A. Meyer and N. Schiller, 319-55. Berlin: Mouton de Gruyter, 2003.

Flege, J. "Second Language Speech Learning Theory, Findings, and Problems." In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by Winifred Strange. Timonium, MD: York Press, 1995.

Flege, J., Munro, M., & Skelton, L. "Production of the Word-Final English /T/-/D/ Contrast by Native Speakers of English, Mandarin, and Spanish." *Journal of the Acoustical Society of America* 92 (1992).

Hammond, M. *The Phonology of English*. New York, NY: Oxford University Press, 1999.

Hansen, J. *Acquiring a Non-Native Phonology*. New York, NY: Continuum, 2006.

Li, W. *A Diachronically-Motivated Segmental Phonology of Mandarin Chinese*. New York: Peter Lang, 1999.

Loewen, S., & Reinders, H. *Key Concepts in Second Language Acquisition*. New York, NY: Palgrave Macmillan, 2011.

Major, R. "Transfer in Second Language Phonology." In *Phonology and Second Language Acquisition*, edited by Jette G. Hansen Edwards and Mary L. Zampini. Philadelphia, PA: John Benjamins Publishing Company, 2008.

Naeser, M. "The American Child's Acquisition of Differential Vowel Duration." Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Tech. Rep. 144, 1970.

Nusbaum, H.C., Pisoni, D.B., & Davis, C.K. "Sizing Up the Hoosier Mental Lexicon: Measuring the Familiarity of 20,000 Words." *Research on Speech Perception Progress Report* 10 (1984): 357-76.

Ortega, L. *Understanding Second Language Acquisition*. London, UK: Hodder Education, 2009.

Raphael, L. "Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-Final Consonants in American English." *Journal of the Acoustical Society of America* 51 (1971): 1296-303.

Shen, X., & Lin, L. "A Perceptual Study of Mandarin Tones 2 and 3." *Language and Speech* 34 (1991): 145-56.

Xu, D.B. *Chinese Phonology in Generative Grammar*. San Diego: Academic Press, 2001.

Appendix A Production Study: Minimal Pairs

Word-Final Plosive	*Bilabial*		*Alveolar*		*Velar*	
	voiceless /p/	*voiced* /b/	*voiceless* /t/	*voiced* /d/	*voiceless* /k/	*voiced* /g/
English Spelling	*cap*	*cab*	*bat*	*bad*	*lock*	*log*
Broad Transcription	/kæp/	/kæb/	/bæt/	/bæd/	/lɑk/	/lɑg/
	gap	*gab*	*hat*	*had*	*tuck*	*tug*
	/gæp/	/gæb/	/hæt/	/hæd/	/tʌk/	/tʌg/
	lap	*lab*	*mat*	*mad*	*pick*	*pig*
	/læp/	/læb/	/mæt/	/mæd/	/pɪk/	/pɪg/
	nap	*nab*	*sat*	*sad*	*rick*	*rig*
	/næp/	/næb/	/sæt/	/sæd/	/rɪk/	/rɪg/
	tap	*tab*	*bit*	*bid*		
	/tæp/	/tæb/	/bɪt/	/bɪd/		
	mop	*mob*	*hit*	*hid*		
	/mɑp/	/mɑb/	/hɪt/	/hɪd/		
	cup	*cub*	*kit*	*kid*		
	/kʌp/	/kʌb/	/kɪt/	/kɪd/		
	pup	*pub*	*lit*	*lid*		
	/pʌp/	/pʌb/	/lɪt/	/lɪd/		

Appendix B Perception Study: Modified and Unmodified Minimal Pairs

Word-Final Plosive	*Bilabial*		*Alveolar*		*Velar*	
	voiceless /p/	*voiced* /b/	*voiceless* /t/	*voiced* /d/	*voiceless* /k/	*voiced* /g/
English Spelling	*cap*	*cab*	*bat*	*bad*	*lock*	*log*
Broad Transcription	/kæp/	/kæb/	/bæt/	/bæd/	/lɑk/	/lɑg/
Unmodified	[kæp]	[kæ:b]	[bæt]	[bæ:d]	[lɑk]	[lɑ:g]
Modified	[kæ:p]	[kæb]	[bæ:t]	[bæd]	[lɑ:k]	[lɑg]
	gap	*gab*	*hat*	*had*	*tuck*	*tug*
	/gæp/	/gæb/	/hæt/	/hæd/	/tʌk/	/tʌg/
	[gæp]	[gæ:b]	[hæt]	[hæ:d]	[tʌk]	[tʌ:g]
	[gæ:p]	[gæb]	[hæ:t]	[hæd]	[tʌ:k]	[tʌg]
	lap	*lab*	*mat*	*mad*	*pick*	*pig*
	/læp/	/læb/	/mæt/	/mæd/	/pɪk/	/pɪg/
	[læp]	[læ:b]	/mæt/	/mæd/	[pɪk]	[pɪ:g]
	[læ:p]	[læb]	/mæt/	/mæd/	[pɪ:k]	[pɪg]
	nap	*nab*	*sat*	*sad*	*rick*	*rig*
	/næp/	/næb/	/sæt/	/sæd/	/rɪk/	/rɪg/
	[næp]	[næ:b]	[sæt]	[sæ:d]	[rɪk]	[rɪ:g]
	[næ:p]	[næb]	[sæ:t]	[sæd]	[r:ɪk]	[rɪg]
	tap	*tab*	*bit*	*bid*		
	/tæp/	/tæb/	/bɪt/	/bɪd/		
	[tæp]	[tæ:b]	[bɪt]	[bɪ:d]		
	[tæ:p]	[tæb]	[bɪ:t]	[bɪd]		
	mop	*mob*	*hit*	*hid*		
	/mɑp/	/mɑb/	/hɪt/	/hɪd/		
	[mɑp]	[mɑ:b]	[hɪt]	[hɪ:d]		
	[mɑ:p]	[mɑb]	[hɪ:t]	[hɪd]		
	cup	*cub*	*kit*	*kid*		
	/kʌp/	/kʌb/	/kɪt/	/kɪd/		
	[kʌp]	[kʌ:b]	[kɪt]	[kɪ:d]		
	[kʌ:p]	[kʌb]	[kɪ:t]	[kɪd]		
	pup	*pub*	*lit*	*lid*		
	/pʌp/	/pʌb/	/lɪt/	/lɪd/		
	[pʌp]	[pʌ:b]	[lɪt]	[lɪ:d]		
	[pʌ:p]	[pʌb]	[lɪ:t]	[lɪd]		