English *tr* Produced by Mandarin Speakers

Ya Li

University of Victoria

yali@uvic.ca

ABSTRACT

This pilot study aims to investigate English *tr* produced by Mandarin speakers through an acoustic analysis. In Mandarin, the consonant cluster *tr* is not permitted. However, Mandarin has the aspirated retroflex affricate, /tʂ/ (orthographically *ch* in Mandarin), which is phonetically close to both English *tr* (/tɹ/) and *ch* (/tʃ/) in terms of the place of articulation (all produced in the post-alveolar area). According to Flege's (1995) Speech Learning Model (SLM), these similar sounds may interfere with one another in both second language perception and production. Previous studies also find that linguistic environment such as the vowel context has a significant effect on consonant or consonant cluster production (Hansen, 2001 & 2006).

Based on the above model/finding, this study aims to address the following research questions: 1) How do English *ch* or Mandarin *ch* interfere with Mandarin speakers' production of English *tr*? 2) What role does the vowel context play in the production? Eight Chinese participants were recruited to perform a reading task containing four English monosyllabic words, *true*, *chew*, *tree*, *cheer* and one Mandarin word, *chù* ('place'). Two participants' speech data were further converted into spectrograms and acoustically analyzed based on the formant patterns.

Two major findings emerged from the acoustic analysis: 1) one Mandarin speaker's production of *tr* occurs as *ch* only when the following vowel is /u/, indicating the vowel context is crucial in determining the substitution of *ch* for *tr*; 2) the other Mandarin speaker's production of *tree* is similar to /tʃɥi/, where *ɥ* is a high front rounded glide in Mandarin. Since /ɥ/ shares the lip-rounding feature with /ɹ/, /tʃɥi/ and /tɹi/ are similar in terms of both the place and the manner of articulation. Also, the production of /tɹi/ as /tʃɥi/ turns the English syllable structure CCV into the permissible Mandarin syllable structure CGV. Overall, the findings not only confirm the model/finding mentioned earlier, but also show that substitution between similar sounds can take place at the syllabic level as well as the segmental level.

Keywords: Consonant cluster *tr*; acoustic analysis; substitution at the syllabic level; vowel context.

1 Introduction

The researcher observed that Chinese speakers often produced English *tr* differently from native speakers; for example, *true* is produced like *chu*, similar to a Mandarin syllable, *chù*. The present study will investigate Mandarin speakers' production of the English consonant cluster *tr* through an acoustic analysis. The organization of this paper is as follows: Section 2 will review previous studies in second language (L2) consonant cluster production and provide the theoretical framework for the present study; Section 3 will explain the research methodology and Section 4 will provide the acoustic analysis of speech data and present important findings; Section 5 will discuss the theoretical and pedagogical implications of the findings for second language production; Section 6 will summarize the study and propose directions for future research.

2 Previous studies

2.1 L2 consonant cluster production

Research on L2 consonant cluster production has been conducted from different perspectives. Earlier researchers focused on how differences between the first language (L1) and L2 sound structure may interfere with L2 acquisition (James, 1980). For example, Cantonese speakers tend to simplify English consonant clusters by deleting a consonant, since most Chinese dialects such as Mandarin and Cantonese do not permit consonant clusters (Tarone, 1987).

Later studies also find the effect of linguistic environment on L2 consonant cluster production. For example, Anderson's (1987) study reveals that Mandarin speakers delete *r* mostly in final clusters and word-medial sequence (in Hansen, 2001). Hansen's (2001) study also shows homo-voicing coda clusters favor absence in Mandarin speakers' production. In addition, Benson's (1988) study finds that Vietnamese speakers tend to delete a final consonant when it follows a diphthong vowel (mentioned in Hansen, 2006).

These studies tend to identify the reduction of consonant clusters as the major type of error in L2 production. However, I observed that Chinese speakers often substituted *tr* with the similar sounding affricate *ch* instead of deleting *r* in *tr*. In fact, Flege's (1995) Speech Learning Model (SLM) claims that similarities rather than differences between L1 and L2 segments are the major source of difficulty for L2 speech learners (in Tarone, 2005).

Based on the above model/findings, the present study aims not only to provide acoustic evidence for Flege's (1995) segment-based SLM, but also to show that the production of the consonant cluster *tr* can be largely influenced by the following vowel.

2.2 Research questions

Two research questions are to be addressed in this study: 1) How do English *ch* or Mandarin *ch* interfere with Mandarin speakers' production of English *tr*? 2) What role does the vowel context play in the production?

3 Research methodology

3.1 Participants

Eight Mandarin Chinese speakers (7 males and 1 female), aged between 18-26, all Chinese international students at University of Victoria (UVic), participated in my study. Also, a native English male speaker's production were recorded and used to illustrate a standard way to produce *tr* and *ch*.

3.2 Instruments

3.2.1. Word-list reading task

The word-list contains 4 English and 1 Mandarin monosyllabic words starting with *tr* or *ch* and ending with the high back vowel /u/ or the high front vowel /i/, as shown in Table 1 (see also Appendix A):

Table 1
The test words used in the word-list reading task

C (C) V	/u/	/i/
/tɹ/ *tr*	true	tree
/tʃ/ *ch*	chew	cheer
/tʂ/ *ch*	chù[1]	n/a[2]

In Table 1, *ch* is used to contrast with *tr*; the two vowels, /u, i/, are used to investigate the effect of linguistic environment on the production of *tr*. A Mandarin word, *chù* (with the fourth tone on *u*) is also included for the Chinese participants to produce. In total, 15 (5 words x 3 tokens) tokens are included.

The Mandarin retroflex affricate /tʂ/ (orthographically *ch* in Mandarin) is phonetically close to both English *tr* and English *ch* in terms of the place of articulation, as they are all produced in the post-alveolar area. According to Flege's (1995) SLM, these similar sounds may interfere with one another in the English production of the Mandarin speakers.

1 To avoid orthographical interference, the Chinese word is written in Chinese in the actual reading task sheet.
2 Mandarin does not have *chì*.

3.2.2. Questionnaire

The questionnaire (see Appendix B) is administered to elicit relevant personal data of Chinese participants, such as gender, major, school year, and the length of the English education (see Appendix C for a summary of the data). The information is used to examine possible correlations between the speakers' background and their production.

3.3 Procedures

1) Eight Mandarin participants were recorded by using Olympus Digital voice recorder (WS-100) in a quiet room.
2) After the 8 participants' production was carefully listened to by the researcher, 2 male participants with clear articulation and typical pronunciation were chosen from the 8 participants and further recorded in the phonetic laboratory at UVic, using Audacity 1.2.4. The native English male participant was also recorded in the same phonetic laboratory.
3) The 3 participants' speech data were analyzed, using Praat 4.4.22. Specifically, the waveform and spectrogram of each token produced by each participant were examined. Then a relatively typical spectrogram was chose from 3 tokens to represent each word. In total, 14 spectrograms (5 words x 2 Chinese participants + 4 words x 1 English participant) were collected and compared/contrasted. Eleven of them will be illustrated in the following acoustic analysis (the other three spectrograms can not be used here due to fuzziness).

4 The acoustic analysis of the speech data and findings

4.1 Basic acoustic properties of tr and ch

The analysis is based on the articulatory mechanisms of English *tr*, *ch* and Mandarin *ch*. Generally, *tr* is considered to be a co-articulated consonant cluster in English, while both English *ch* and Mandarin *ch* are affricates. *tr* is a combination of the voiceless stop /t/ and the rhotic approximant /ɹ/; English *ch*, on the other hand, is a combination of the voiceless stop /t/ and a voiceless fricative /ʃ/. Similarly, Mandarin *ch* is a combination of the voiceless stop /t/ and a voiceless fricative /s̺/. In fact, Ladefoged and Maddieson (1996) classify Mandarin /s̺/ as the flat post-alveolar fricative rather than a retroflex. In contrast, English /ʃ/ is classified as the domed post-alveolar fricative. Therefore, English *ch* and Mandarin *ch* are very similar in terms of not only the manner of articulation but also the place of articulation, because they both can be considered as post-alveolar affricates.

Figure 1, the spectrograms of *true* and *chew*, and Figure 2, the spectrograms of *tree* and *cheer* are extracted from the native English speaker's production. The difference between productions of *tr* and *ch* can be observed clearly in the four spectrograms. The first segment /t/ in *tr* and *ch* is represented by a blank area followed by a vertical line (see the spectrograms of *chew* and *cheer*). The blank area indicates the silent period when articulators (the tongue tip and the alveolar) are closed; the vertical line indicates the release burst of the closed articulators (Hayward, 2000). The second element /ɹ/ in *tr* comprises two parts: the voiceless and the voiced; that is, /ɹ/ in *tr* is partially devoiced to assimilate the preceding voiceless stop /t/. Notice the two parts together form a V-shaped pattern (see the spectrogram of *tree*). The voiceless part of /ɹ/ (phonetic representation, [ɹ̥]) can be observed as vertical striations immediately following the vertical line, indicating noisy airflow that accompanies voicelessness. The voiced part of /ɹ/ includes a rising F3 (the third formant), indicating the F3 of /ɹ/ is lower than the following vowel (see the spectrogram of *true*). Note in the spectrogram of *tree*, the lower part of F3 is conflated with F2.

Unlike /ɹ/, the second element /ʃ/ in English *ch* only has striations across the upper frequency range, indicating frictions caused by airflow passing through narrowed articulators (Davenport & Hannahs, 1998). Notice the onset part of *tr* (representing /tɹ̥/) is similar to /tʃ/, both with a vertical line followed by massive striations (see Figure 1). Also the spectrogram of *chew* shows a palatal sound /j/ between /ʃ/ and /u/, a typical pronunciation for some English speakers.

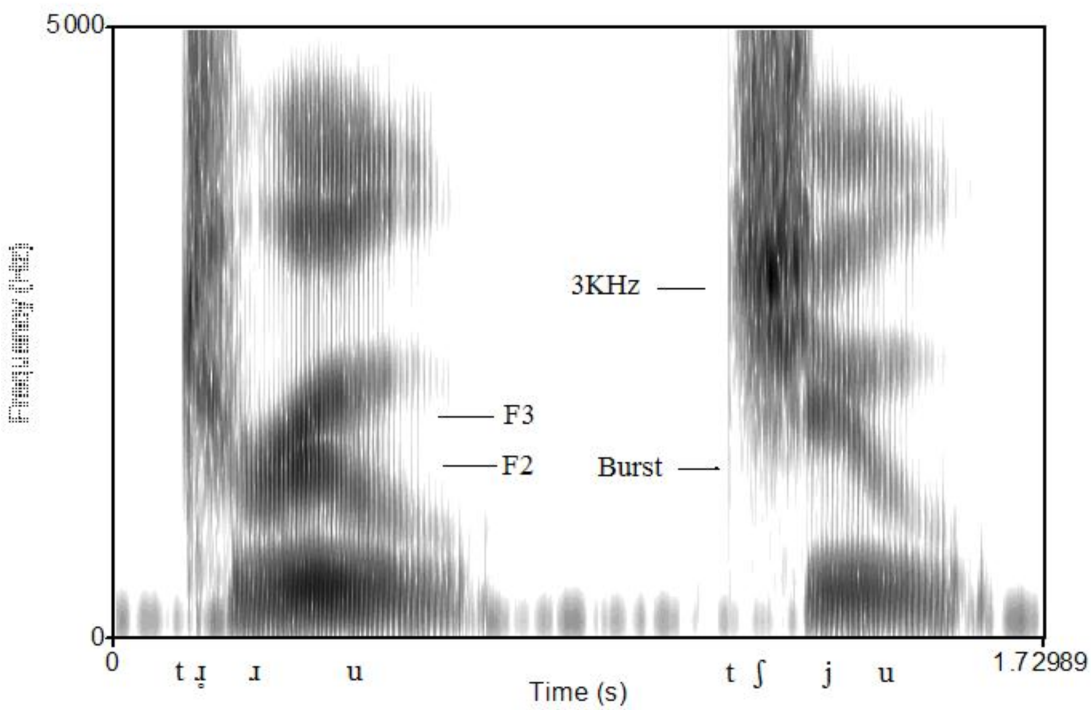

Figure 1
Spectrograms of *true* (left) and *chew* (right) produced by the native English speaker.

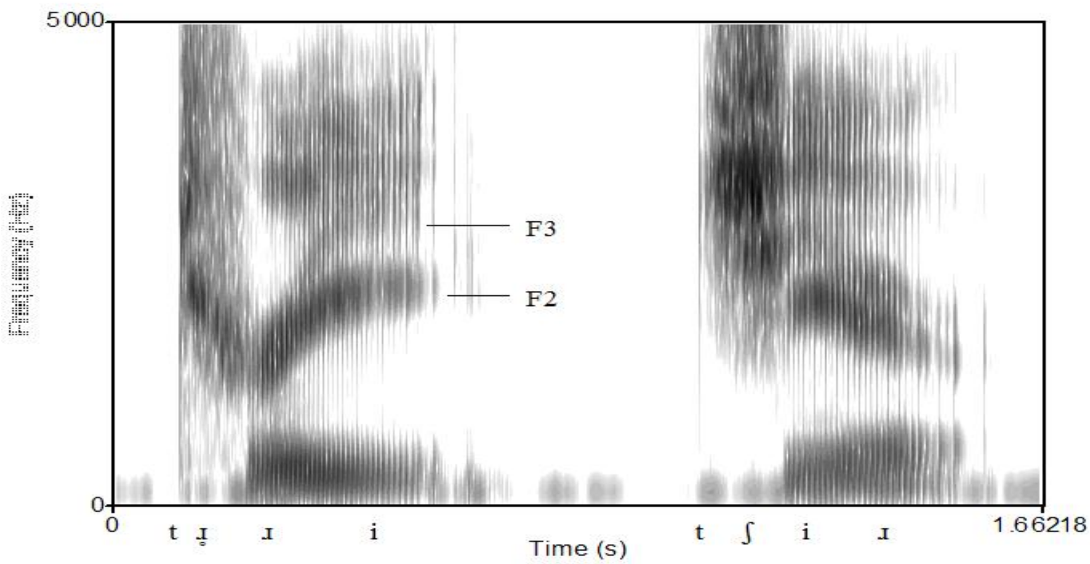

Figure 2
Spectrograms of *tree* (left) and *cheer* (right) produced by the native English speaker

Figure 3 provides the spectrograms of *true* and *chew* produced by the first Mandarin male speaker. The two spectrograms show almost identical onset patterns: a vertical line (the release burst for /t/) precedes striations. Notice the V-shaped formant pattern for /ɹ/ is missing from the spectrogram of the presumably *true* (i.e., The speaker thought he was producing *true)*, indicating that this speaker produced *tr* similar to *ch* when the following vowel is /u/.

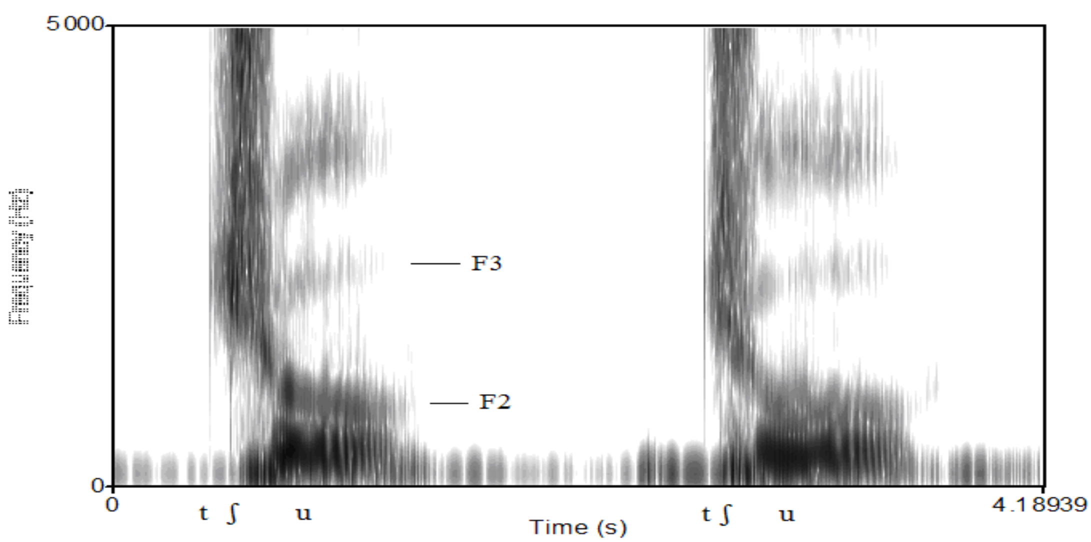

Figure 3
Spectrograms of *true* (left) and *chew* (right) produced by the first Mandarin speaker

Figure 4 shows how the same speaker produced *ch* in the Mandarin word *chù*. Similar to the spectrograms of *true* and *chew* in Figure 3, the onset part of *chù* has massive striations, and no V-shaped formant pattern is present, so it seems hard to differentiate this speaker's productions of English *tr, ch,* and Mandarin *ch* when the following vowel is /u/. Sometimes, based on the frequencies around the darkest striations, it is possible to infer which fricative is involved. According to Davenport & Hannahs (1998), the higher the frequencies, the further forward the place of articulation is. Thus, the post-alveolar fricative /ʃ/ has the frequencies around 3kHz, the retroflex fricative /ʂ/ around 2kHz. For example, In Figure 1, the native speaker's production of /ʃ/ has the darkest striations around 3KHz. For this Mandarin speaker, however, it is hard to identify the height of the darkest striations in the spectrograms of *true*, *chew*, and *chù* (see Figures 3 & 4). The darkest striations for *true* and *chew* start almost at the same height around 2KHz, again indicating *tr* was similarly produced as *ch*.

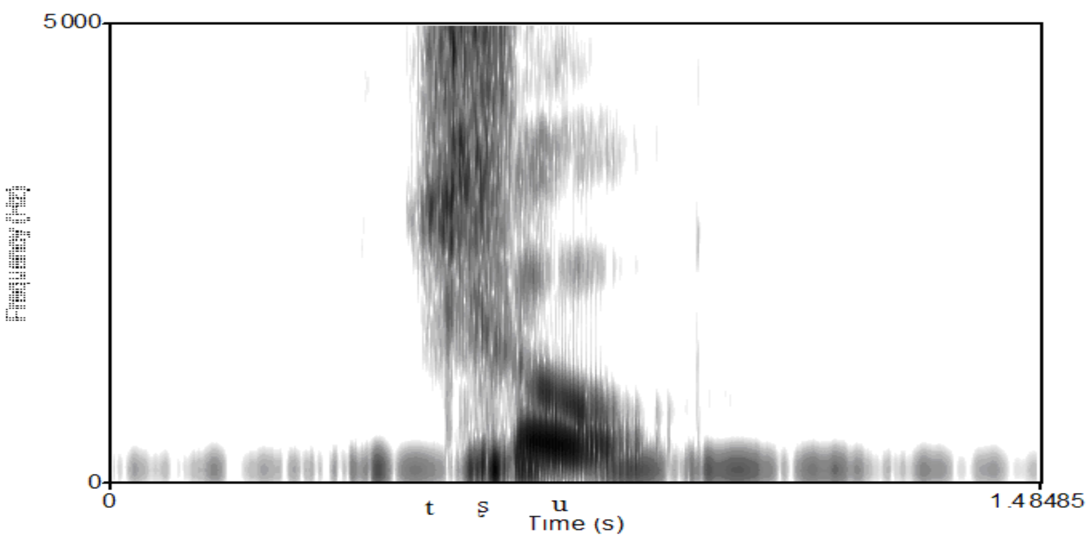

Figure 4
Spectrogram of the Mandarin word *chù* produced by the first Mandarin speaker

In Figure 5, however, there is a noticeable difference between the speaker's productions of *tree* and *cheer*. In the spectrogram of *tree*, F2 and F3 gradually move upward following the striations at the onset, which forms a rather narrow V-shape (indicating /ɹ/). In the spectrogram of *cheer*, only the high frequency striations (indicating [ʃ]) are present, followed by the downward-going F2 and F3 signaling transitions from /i/ to the final /ɹ/ in *cheer*. The difference between the onset patterns of the two spectrograms indicates that the speaker was able to treat *tr* and *ch* differently when the following vowel is /i/.

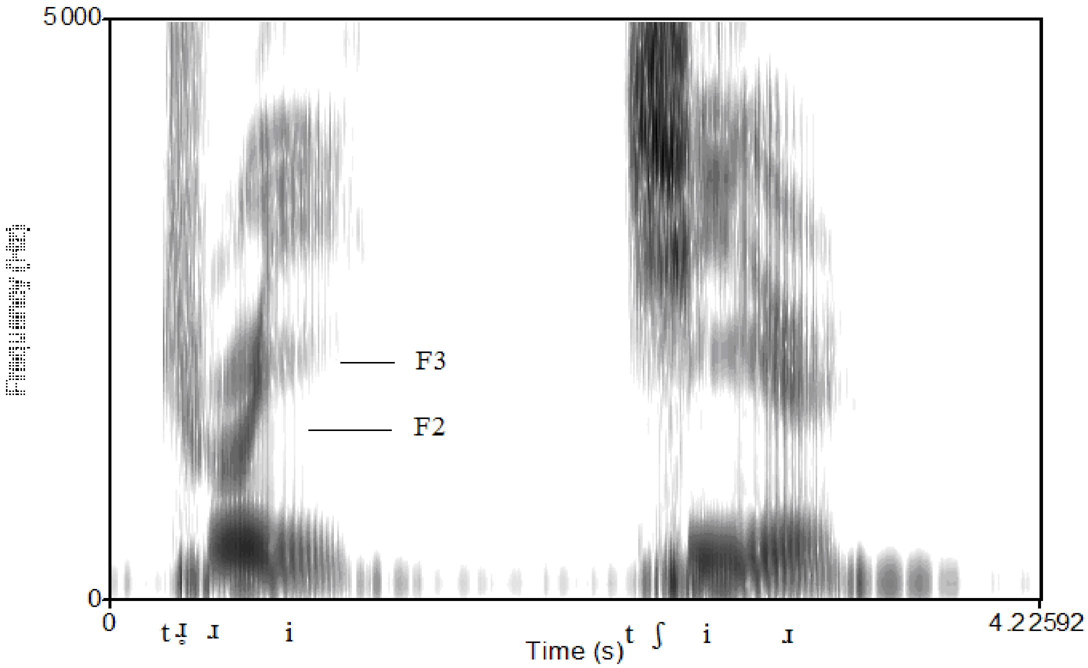

Figure 5
Spectrograms of *tree* (left) and *cheer* (right) produced by the first Mandarin speaker

Figure 6 illustrates how the second Chinese male speaker produced *tree* and *cheer*. In contrast to the first Chinese speaker, this speaker's production of *tree* is similar to /tʃɥi/, with a lip-rounding glide /ɥ/ in between the affricate /tʃ/ and the vowel /i/. In the spectrogram of the presumably *tree*, the striations are followed by gradually rising F2, F3, and F4. According to Fant (2004), lip rounding has a general effect of lowering formants, so the rising formant transitions indicate the change of place of articulation from a high front rounded position towards a high front unrounded position (i.e., from the glide /ɥ/ to the target vowel /i/). Notice that although the production of /ɹ/ also involves lowering F2 and F3 (i.e., /ɹ/ and /ɥ/ both share the lip-rounding feature), the formant pattern of /ɹ/ can be still distinguished from that of /ɥ/ by the rapidly lowering F3 and the absence of the higher formants such as F4 and F5.

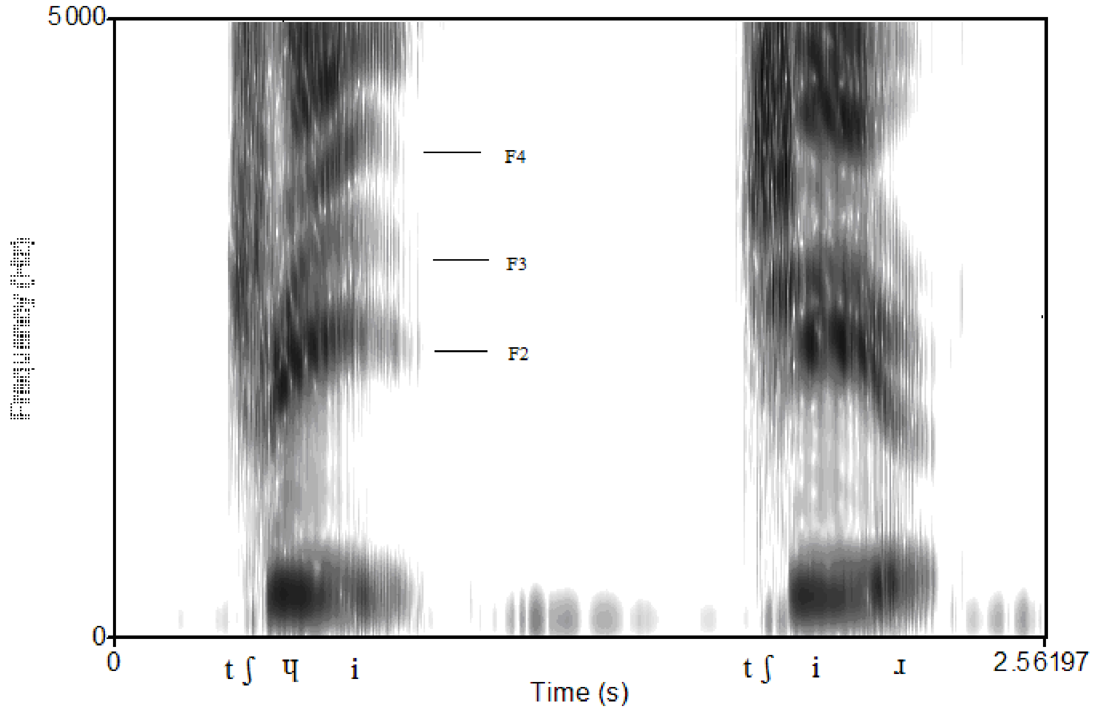

Figure 6
Spectrograms of *tree* (left) and *cheer* (right) produced by the second Mandarin speaker

Based on the results from the above analysis, the two research questions can be answered. 1) How do English *ch* or Mandarin *ch* interfere with Mandarin speakers' production of English *tr*? Figure 3 shows that the production of English *tr* by the first speaker occurs as *ch* only when the following vowel is /u/. Note that the distinction between English *ch* and Mandarin *ch* is not clear in this speaker's production. 2) What role does the vowel context play in the production? Figure 5 shows that the production of *tr* by the first speaker is more or less native-like when the following vowel is /i/, but Figure 6 shows that the production of *tr* by the second speaker is influenced by *ch* in the same vowel context. Note the influence involves substituting /tɹi/ with the syllable /tʃɥi/ rather than /tʃi/ alone.

5 Theoretical and pedagogical implications

The above results also raise the following question: why can the first Mandarin speaker produce *tr* native like in *tree* but not in *true*? An explanation could be that /tɹu/ and /tʃu/, both sharing the acoustically prominent lip rounding feature (i.e., the rounded /u/), are similar enough for substitution; while that /tʃi/, a potential candidate for substitution, does not share the lip rounding feature with /tɹi/, so they are not similar enough for substitution. As a result, the speaker may have learned to produce *tree* native-like.

However, if /ɥ/ is inserted into /tʃi/, then /tʃɥi/ and /tɹi/ both share the lip-rounding feature (/ɥ/ and /ɹ/ both involve lowered F3) and therefore become very similar. In other words, when the glide /ɥ/ is used to fill the missing lip-rounding feature in /tʃi/, the substitution of /tʃ/ for /tɹ/ becomes possible, as seen in the second speaker's production of *tree*. Also, the production of /tɹi/ as /tʃɥi/ turns the English syllable structure CCV into the permissible Mandarin syllable structure CGV. In fact, after the researcher carefully listened to the other 6 Chinese participants' production, /tʃɥi/[3] is found to be a preferred production instead of /tɹi/ (clearly occurred in 4 out of 8 speakers' production). Since the substitution of /tʃɥi/ for /tɹi/ takes place at both the syllabic level and the segmental level, Flege's (1995) SLM model can be expanded to include similarities between syllables, not just segments, as a source of interference.

Pedagogically, the above findings have two implications. First, the traditional minimal pair drills can include the consonant cluster perception and production. For example, the words used in the reading task such as *true* and *chew* can be used to train Chinese ESL learners to differentiate *tr* and *ch* in different vowel contexts. Second, ESL teachers are necessarily trained to know the phonological structure (including syllable structure) of their students' first language. Suppose an English teacher knows how the Mandarin *chù* is produced. He or she can include it in the minimal pair training with *true* and *chew*. This method is in fact the essence of "contrastive teaching" discussed by James (1988). The purpose of contrastive teaching is to make the students put more conscious effort in distinguishing between their native sounds and the sounds of their second language (James, 1988). For example, after the researcher explained in plain language to a participant how *true* was different from and similar to English *chew* and Mandarin *chù* in terms of the articulatory mechanism and the syllable structure, his ability to distinguish these words has improved in an informal perception test conducted by the researcher. The improved perception in turn might be able to help to improve this participant's English production.

6 Conclusions

Although the data in this pilot study are relatively scanty, the results are informative: interference between the consonant cluster *tr* and the affricate *ch* is based mainly on their similar features such as the post-alveolar articulation and the voicelessness after the release of *t*. However, the interference is caused not only by the similarities between the two sounds, but also determined by the following vowel. Depending on the vowel context, different substitutes for *tr* occurred in the two Mandarin speakers' production; that is, /tʃu/ was used to replace /tɹu/ and /tʃɥi/ was used to replace /tɹi/. These results provide evidence to both support and extend Flege (1995)'s SLM model, suggesting that interference based on similarities between two sounds can take place not only at the segmental level but also at the syllabic level.

---

[3] Although /tʃɥi/ conforms to the Mandarin CGV structure, it is not a Mandarin syllable.

This study examines only eight Mandarin speakers' production of *tr*, so more participants can be included in future research. Also, more test words can be chosen to measure further influence of linguistic environment. For example, future research could compare *tree* with the Mandarin words *qì* and *qù*, where *q* (/tɕ/) is a palatal affricate similar to *ch* (/tʃ/) and *ù* in *qù* is a high front rounded vowel corresponding to the high front rounded glide /ɥ/ in Mandarin. In a word, based on the present study, future research can be conducted on a larger scale.

Acknowledgments

I would like to thank Dr. Li-Shih Huang for her helpful comments and my fellow graduate student Thomas Magnuson for providing the native production data in the study. Also, I would like to thank my supervisor Dr. Hua Lin for encouraging me to submit the abstract to the 23th NWLC conference.

Appendix A. Reading Task

1. Please read the following English words slowly.

a.	cheer	cheer	cheer
b.	chew	chew	chew
c.	true	true	true
c.	tree	tree	tree

2. Please read the following Mandarin words slowly (for Mandarin speakers).

处 处 处

(处: /chù/, 'place')

Appendix B. Questionnaire

Thank you for participating in my language survey about the English pronunciation of Chinese students. There are no right or wrong answers. The data will be used to provide background information for my study. No identification information will be asked in this survey. The consent form will be detached from both this page and from the recording.

Gender: School year: Major:

1. Do you speak a Chinese dialect other than Mandarin? If so, please specify the dialect you speak, for example, "Beijinghua" and "Cantonese."

2. At what age did you start to learn English?

3. Have you taken or are taking any ESL courses in Canada or other English speaking countries?

Appendix C. Summary of 8 Chinese participants' personal data

The total number of participants in this study	8 (1 female and 7 males)
Age range	From 18 to 26
School year range	6 from 1st year to 4th year, 2 in MBA 1st year
Major (number of participants)	Economics (5), MBA (2), Electronic Engineering (1)
Chinese dialects other than Mandarin (number of participants)	Beijing (1), Taiwanese (1), Hubei (1), Nanchang (3), Cantonese (1), Taizhou (1)
Age starting to learn English (number of participants)	3 (1), 4 (1), 11 (4), 12 (1), 14 (1)
Number of participants who took ESL courses (location)	7 (in Canada), 1 (in USA)

References:

Anderson, J. (1987). The markedness differential hypothesis and syllable structure difficulty. In Ioup, G. and Weinberger, S.H. (eds.). *Interlanguage phonology: The acquisition of a second language sound system* (pp.279-291). Cambridge, MA: Newbury House Publishers.

Benson, B. (1988). Universal preference for the open syllables as an independent process in interlannguage phonology. *Language Learning*, 38 (2): 221-235.

Davenport, M. & Hannahs, S.J. (1998). *Introducing phonetics and phonology*. London: Arnold.

Fant, G. (2004). *Speech acoustics and phonetics*. Dordrecht: Kluwer Academic Publishers.

Flege, J. (1995). Speech learning in a second language. In Ferguson, C., Menn, L. and Stoel-Gammon, C. (eds.). *Phonological development* (pp.565-604). Timonium, MD: York Press.

Hansen, J.G. (2001). Lingusitic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Lingusitics* 22 (3): 338-365.

Hansen, J. G. (2006). *Acquiring a non-native phonology: linguistic constraints and social barriers*. London. Continuum.

Hayward, K. (2000). *Experimental phonetics*. Harlow: Longman.

James, C. (1980). *Contrastive Analysis*. Harlow: Longman.

James, C. (1988). *The acquisition of a second language phonology: A linguistic theory of developing sound structures*. Tübingen: Narr.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Cambridge: Blackwell.

Tarone, E. (1987). Some influences on the syllable structure of interlanguage phonology. In Ioup, G. and Weinberger, S.H. (eds.). *Interlanguage phonology: The acquisition of a second language sound system* (pp.233-247). Cambridge, MA: Newbury House Publishers.

Tarone, E. (2005). Speaking in a second language. In Hankel, Ed. *Handbook on research in second language teaching and learning* (pp.485-502). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers.