



**164th Meeting of the Acoustical Society of America  
Kansas City, Missouri  
22 - 26 October 2012**

**Session 2aSC: Speech Communication**

**2aSC16. American Chinese learners' acquisition of L2 Chinese affricates /ts/ and /tsh/**

**Jiang Liu\* and Allard Jongman**

**\*Corresponding author's address: Linguistics, The University of Kansas, Lawrence, KS 66045, [liujiang@ku.edu](mailto:liujiang@ku.edu)**

Many studies on L2 speech learning focused on testing the L1 transfer hypothesis. In general, L2 phonemes were found to be merged with similar L1 phoneme to different degrees (Flege 1995). Few studies examined whether non-phonemic phonetic categories in L1 help or block the formation of new phonetic categories in L2. The current study examined the effect of L1 English consonantal clusters [ts] and [dz] on learning L2 Chinese affricates /ts/ and /tsh/. We studied duration and center of gravity (COG) of Chinese affricates /ts/ and /tsh/ produced by native Chinese speakers, novice American Chinese learners and advanced learners. In terms of duration, both learner groups showed contrast between L2 /ts/ and /tsh/, which is similar to native Chinese speakers' production. However, for COG, only the advanced learner group showed contrast between L2 /ts/ and /tsh/, which is similar to native speakers' production while the novice learner group did not show a COG difference between the two L2 affricates. The results suggest an early acquisition of the durational contrast between the L2 Chinese affricates and later acquisition of COG contrast between the two L2 affricates.

Published by the Acoustical Society of America through the American Institute of Physics

## 1. Introduction

In the field of second language speech learning, many studies have been conducted to examine the nativelikeness of L2 production and perception. In terms of production, the nativelikeness examined in most studies is mainly evaluated based on the comparison between the acoustic characteristics of the phonemes produced by L2 learners and native speakers. In terms of perception, the nativelikeness examined in most studies is mainly evaluated based on whether L2 learners use the same acoustic cues as native speakers to identify different phonemes, especially, based on the comparison between L2 learners and native speakers' weightings of different acoustic cues (Strange 1995). The majority of the studies on L2 speech learning focus on testing the L1 transfer hypothesis. The merging of L1 and L2 phonemes is accounted for due to the similarity between the phonemes in the two languages. Thus, the L1 transfer hypothesis is generally supported (Best 1993, Flege 1995). Few studies have examined whether non-phonemic phonetic categories in L1 help or block the formation of new phonetic categories in L2. The current study examines the effect of L1 English consonantal clusters [ts] (e.g., the ending of the plural noun 'fruits') and [dz] (e.g., the ending of the plural noun 'foods') on the acquisition of the L2 Chinese affricates /ts/ and /ts<sup>h</sup>/. While English [ts] and [dz] are consonantal clusters with a fricative following a stop, the Chinese affricates /ts/ and /ts<sup>h</sup>/ are single consonants. Impressionistically, English [ts] and [dz] sound similarly to Chinese /ts<sup>h</sup>/ and /ts/. Thus, even though [ts] and [dz] are not phonemes in English, the fact that they occur as clusters may affect the acquisition of the L2 Chinese affricates. Since the extent of L2 Chinese experience may affect the learners' production of L2 Chinese /ts/ and /ts<sup>h</sup>/ we examined the production of the two L2 Chinese affricates by learners with different levels of L2 Chinese proficiency.

Although English consonantal clusters [ts] and [dz] may sound similar to Chinese affricates /ts/ and /ts<sup>h</sup>/, they differ phonotactically. In Chinese, the affricates /ts/ and /ts<sup>h</sup>/ can only occur in syllable onset position whereas in English the consonantal clusters [ts] and [dz] never occur in syllable onset position. Therefore, the potential difficulty for learning the L2 Chinese affricates /ts/ and /ts<sup>h</sup>/ lies not only in learning the new phonemes but also producing them in syllable-initial position. On the basis of the perceptual similarity and the phonotactic difference between the English consonantal clusters and the Chinese affricates, the current study addresses the following two research questions:

1. Do American learners produce a distinction between L2 Chinese /ts<sup>h</sup>/ and /ts/ in terms of temporal and/or spectral properties?
2. Do more proficient learners more closely approximate the native speakers?

There are only a few acoustic studies of Chinese affricates. Zhang and Qi (1982) measured the duration of the release portion (burst plus frication) of Chinese /ts/ and /ts<sup>h</sup>/ and found that aspirated /ts<sup>h</sup>/ was significantly longer than unaspirated /ts/. Averaged across 12 native speakers of Mandarin Chinese, /ts<sup>h</sup>/ had a mean duration of 154ms whereas /ts/ had a mean duration of 44ms. In spectral terms, aspiration in fricatives has been studied in languages like Korean (e.g., Lee 2011). In her study of native Korean speakers' productions of Korean aspirated /s<sup>h</sup>/ and unaspirated /s/, aspirated /s<sup>h</sup>/ had a

significantly lower center of gravity than unaspirated /s/. Based on this effect of aspiration on the center of gravity for Korean fricatives, we expect a similar influence of aspiration on Chinese affricates. In the present study, acquisition of the Chinese affricates will be assessed in terms of the temporal acoustic parameter of release duration and the spectral acoustic parameter of center of gravity (see 2.4 for more details).

## 2. Methods

### 2.1 Chinese stimuli

In the production experiment, American learners of L2 Chinese were asked to read a wordlist consisting of 16 pairs of Chinese monosyllables starting with /ts/ or /ts<sup>h</sup>/. The tones and rhymes of the monosyllables were controlled. The target syllables in the wordlist were shown to the subjects in both Chinese characters and their corresponding pinyin<sup>1</sup>. A couple of examples of the word pairs are shown in Table 1.

**Table 1. Examples of Chinese monosyllables starting with /ts/ and /ts<sup>h</sup>/**

/ts/	zāi 裁	zū 租
/ts <sup>h</sup> /	cāi 猜	cū 粗

### 2.2 English stimuli

A wordlist that consisted of 16 pairs of L1 English monosyllables ending with the consonantal clusters [ts] and [dz] was designed. All the monosyllables are either plural forms of nouns or third person singular forms of verbs. The rhymes of the monosyllables were controlled. Examples of the word pairs include ‘shoots’ and ‘foods’.

### 2.3 Participants

Five level-1 American learners of Chinese (learning Chinese for less than 6 months) and five level-3 American learners of English (learning Chinese for about 2.5 years) recorded both the Chinese and English stimuli embedded in carrier sentences. Five native Chinese speakers also recorded the Chinese stimuli.

### 2.4 Recording and acoustic analysis

<sup>1</sup> ‘pinyin’ is a Romanized pronunciation transcription system developed for Chinese characters.

All speakers were recorded in an anechoic chamber at the University of Kansas, using a Marantz PMD-671 digital recorder and an Electro-Voice 767a microphone.

Based on previous acoustic studies, we measured duration of the release portion (burst+frication) and center of gravity of the frication part of Chinese /ts/ and /ts<sup>h</sup>/. Center of gravity (henceforth COG) was measured at 25%, 50% and 75% of the frication part. In general, the aspirated affricate is expected to have a longer release portion and a lower COG than its unaspirated counterpart.

### 3. Results

In terms of duration, we first conducted a 2x3 mixed ANOVA [within-subject factor: Consonant (/ts<sup>h</sup>/ and /ts/); between-subject factor: Proficiency (novice learners, advanced learners and native speakers)]. The mixed ANOVA shows a significant main effect of Consonant ( $F(1,12)=312$ ,  $p<.001$ ) but no main effect of Proficiency. In other words, the the release part of Chinese /ts<sup>h</sup>/ averaged across proficiency levels is significantly longer than that of Chinese /ts/ (/ts<sup>h</sup>: 143ms vs. /ts/: 93ms). However, there is no durational difference among the three proficiency levels averaged across the two Chinese affricates. There is a significant interaction between Consonant and Proficiency ( $F(2,12)=7.7$ ,  $p<.01$ ). Figure 1 illustrates the interaction between Consonant and Proficiency.

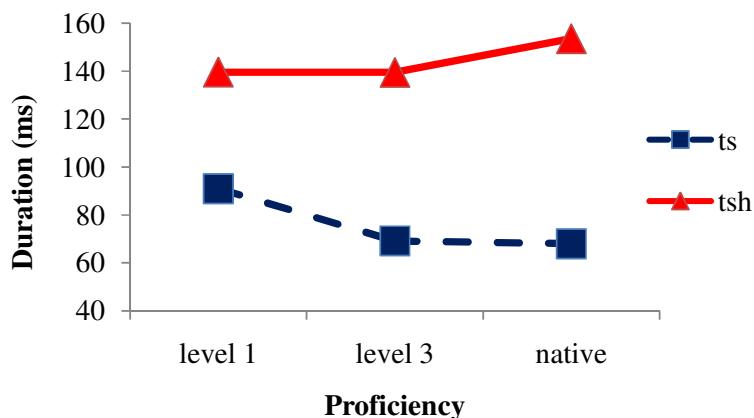


Figure 1. Duration of Mandarin /ts/ and /ts<sup>h</sup>/ as a function of proficiency

Figure 1 shows that the aspirated affricate /ts<sup>h</sup>/ is consistently longer than the unaspirated affricate /ts/; however, the durational difference between the two affricates is smaller for level-1 learners than for the other two proficiency groups.

To examine the nativelikeness of L2 Chinese affricates produced by the learner groups in terms of duration, we conducted a One-way ANOVA to compare the durational difference between L2 Chinese affricates /ts/ and /ts<sup>h</sup>/ across the three speaker groups. The result showed that there was a significant difference among the three speaker groups ( $F(2, 12)=7.74, p<.01$ ). Bonferroni post-hoc tests showed that the level-1 learner group had a significantly smaller durational difference between the two L2 Chinese affricates than the level-3 learner group and the native speaker group. The results are illustrated in Figure 2.

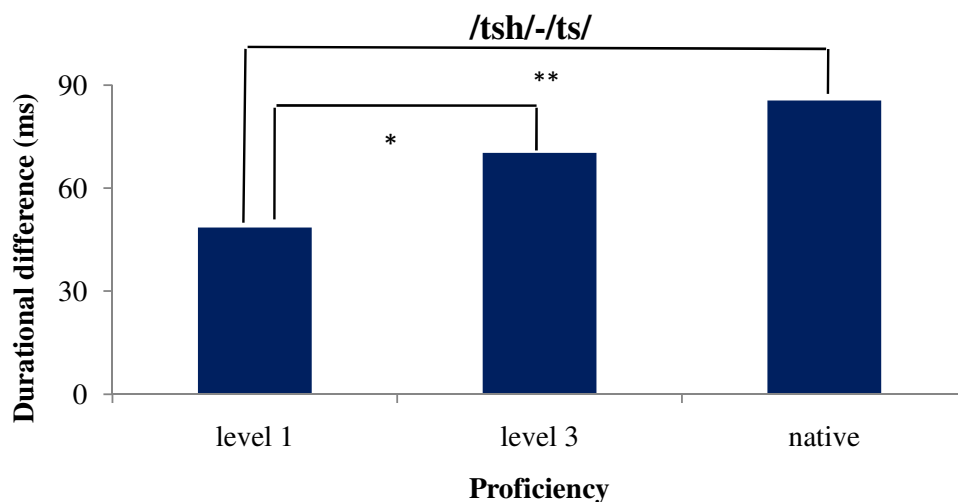


Figure 2. Mean durational difference between the Chinese affricates /ts/ and /ts<sup>h</sup>/ as a function of Proficiency. \* $<.05$ , \*\* $<.01$ .

Level-3 learners behaved like native speakers in terms of the durational contrast between the L2 Chinese affricates /ts/ and /ts<sup>h</sup>/ whereas level-1 learners did not achieve nativelikeness even though they showed a significant durational difference in the right direction between Chinese /ts<sup>h</sup>/ and /ts/.

The center of gravity (COG) was measured at three time points during the frication part of the Chinese affricates /ts<sup>h</sup>/ and /ts/. For the first set of statistical analyses, we use COG averaged across 25%, 50% and 75% of each token as the dependent variable. The averaged COG provides a more stable estimate of the spectral mean. Comparing the averaged COG across three speaker groups can tell us whether the different proficiency groups produce the Chinese affricates differently in terms of overall spectral characteristics. For the second set of statistical analyses, we use the difference in COG at 25% and 75% of each token as the dependent variable to examine whether the three

proficiency groups produce the Chinese affricates differently in terms of dynamic spectral characteristics.

A 2x3 mixed ANOVA (within-subject factor—Consonant: Chinese affricates /ts<sup>h</sup>/ vs. /ts/; between-subject factor—Proficiency: novice learner; advanced learner and native speaker) showed a significant main effect of Consonant [/ts/: 7753Hz vs. /ts<sup>h</sup>/: 6992Hz, (F(1,12)=69.98, p<.001)] but no main effect of Proficiency. The result suggests that the two Chinese affricates differ in terms of average COG. There is a significant interaction between Consonant and Proficiency (F(2,12)=13.64, p<.01) as shown in Figure 3.

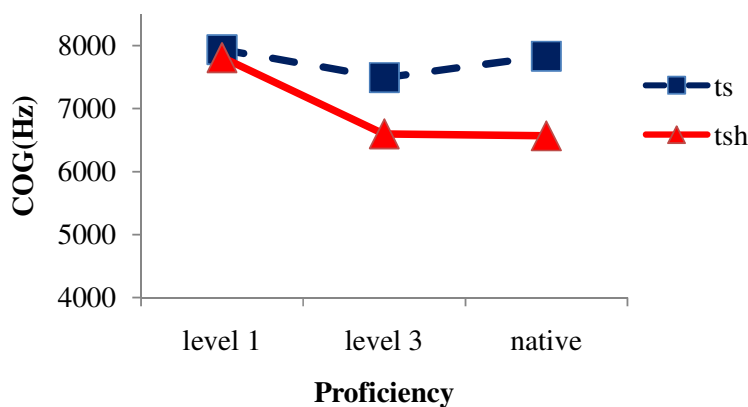


Figure 3. Interaction between consonant and proficiency in terms of COG

The interaction indicates that the COG difference between /ts/ and /ts<sup>h</sup>/ is different across the three speaker groups. Level-3 learners and native speakers had a much larger COG difference between the two Chinese affricates than level-1 learners.

In order to examine the nativelikeness of the L2 Chinese affricates produced by the learner groups in terms of COG, we compared the COG difference between the Chinese affricates /ts/ and /ts<sup>h</sup>/ among the three proficiency groups. Such a comparison can tell us whether the learner groups have a COG contrast between L2 Chinese affricates that is similar to the COG contrast made by the native speakers. A one-way ANOVA shows that COG differences between Chinese affricates are different among the three speaker groups (F(2,12)=13.64, p<.01). Post-hoc tests show that the COG difference between Chinese affricates in the level-1 learner group is significantly smaller than in both the native speaker group and level-3 learner group. The result is illustrated in Figure 4.

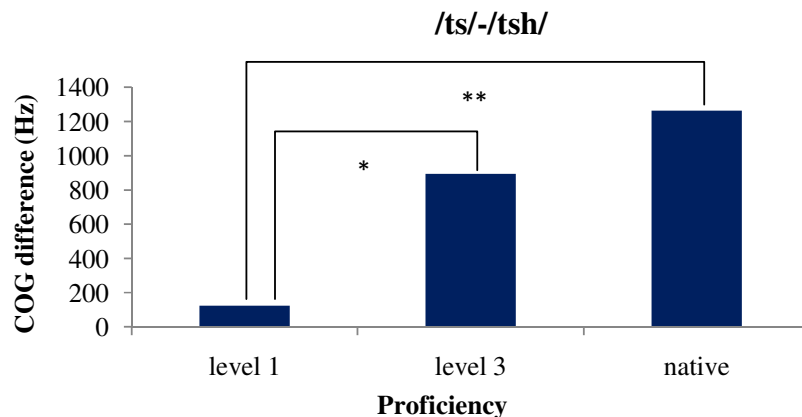


Figure 4. Mean COG difference between the Chinese affricates /ts/ and /ts<sup>h</sup>/ as a function of proficiency. \* $<.05$ , \*\* $<.01$ .

These results suggest that the level-3 but not the level-1 learner group achieved nativelikeness in terms of the COG contrast between the L2 Chinese affricates /ts/ and /ts<sup>h</sup>/.

In terms of dynamic spectral characteristics, we compared COG at 25% and 75% of the frication portion of the affricate. We conducted three 2x2 repeated measures ANOVAs for the three speaker groups, respectively, by adding a new factor, Time Window, with two levels--25% and 75%. This set of ANOVAs can inform us whether the two learner groups have a similar COG lowering from the unaspirated portion to the aspirated portion in the Chinese affricate /ts<sup>h</sup>/ (cf. Lee 2011). The results indicate that all three speaker groups showed a significant main effect of Time Window—level-1 learners (COG at 25%: 8010Hz; COG at 75%: 7634Hz,  $F(1,4)=15.79$ ,  $p<.05$ ); level-3 learners (COG at 25%: 7571Hz; COG at 75%: 6536Hz,  $F(1,4)=58.33$ ,  $p<.01$ ) and Chinese native speakers (COG at 25%: 7839Hz; COG at 75%: 6427Hz,  $F(1,4)=47$ ,  $p<.01$ ). Even though a main effect of Time Window exists for all three speaker groups, the Time Window's influence on COG differs between the level-1 learner group and the other two speaker groups in terms of COG drop for the aspirated affricate /ts<sup>h</sup>. The difference is reflected in the significant interaction between Consonant and Time Window in the three speaker groups. The level-1 learner group did not show such a significant interaction, whereas the other two groups did show a significant interaction between Consonant and Time Window (level-3 learner group:  $F(1,4)=9.7$ ,  $p<.05$ ; native speaker group:  $F(1,4)=231$ ,  $p<.001$ ). The interaction results in the three speaker groups are illustrated in Figure 5.

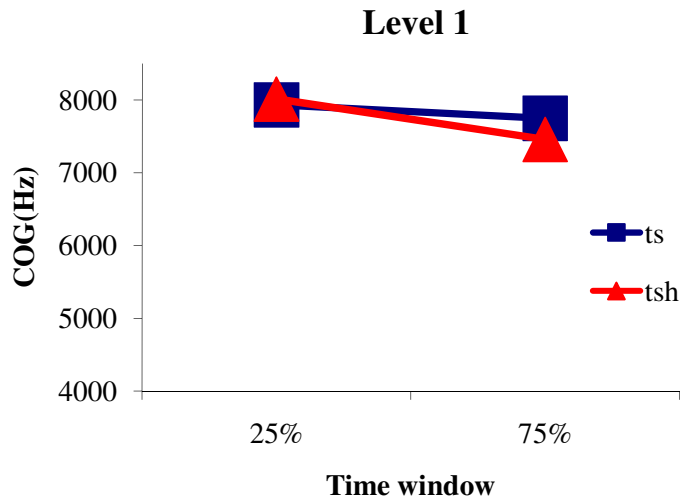


Figure 5a. COG of affricates /ts/ and /ts<sup>h</sup>/ as a function of time window for level-1 learners.

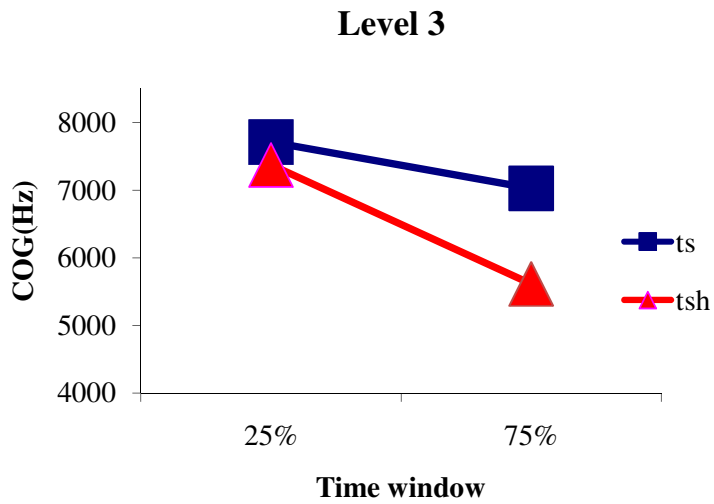


Figure 5b. COG of affricates /ts/ and /ts<sup>h</sup>/ as a function of time window for level-3 learners.



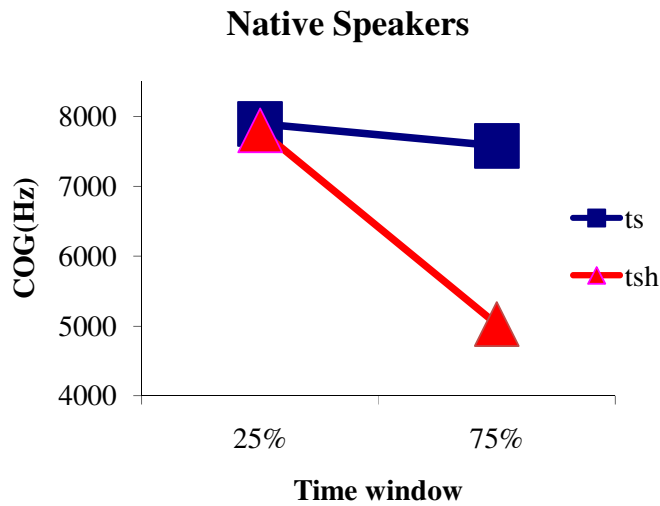


Figure 5c. COG of affricates /ts/ and /ts<sup>h</sup>/ as a function of time window for native Chinese speakers.

The significant interaction between Consonant and Time Window in the level-3 learner group and native speaker group suggests that COG became much lower at the 75% time window relative to the 25% time window for the aspirated affricate /ts<sup>h</sup>/ but not for the unaspirated affricate /ts/. The lack of a significant interaction between Consonant and Time Window for the level-1 learners suggests that any change in COG from the 25% to the 75% time window was similar for both the aspirated /ts<sup>h</sup>/ and unaspirated /ts/. Based on the significant COG drop that occurred in the aspirated /ts<sup>h</sup>/ produced by level 3 learners, we argue that only level 3 learners have acquired the aspiration feature for L2 Chinese /ts<sup>h</sup>/.

#### 4. Conclusion

In the current study, we investigated both temporal and spectral properties of the Chinese affricates /ts/ and /ts<sup>h</sup>/ produced by American learners of Chinese. We found that novice learners did acquire the durational contrast between the Chinese affricates /ts/ and /ts<sup>h</sup>/ but failed to show a center of gravity contrast between the two Chinese affricates. The advanced learners acquired both durational and center of gravity contrasts between the Chinese affricates /ts/ and /ts<sup>h</sup>. The early acquisition of the durational contrast between the two L2 Chinese affricates may be due to the salience of the duration cue for L2 learners. The advanced learners showed both durational and center of gravity contrasts between the L2 Chinese /ts/ and /ts<sup>h</sup>. This suggests a later acquisition of the spectral contrast for the L2 Chinese affricates. This suggests that the acquisition of the aspiration feature together with the affricate takes time, but is possible

## References

Best, C. (1993) Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, ed. B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, and J. Morton. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Flege, J. (1995) Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. W. Strange, Baltimore: York Press. pp 324-358.

Lee, G. (2011) *Acoustic characteristics of Korean fricatives and affricates*. MA thesis, The University of Kansas.

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

Zhang, J. L. & Qi, S. (1982) A cluster analysis of the perceptual features of Chinese speech sounds. *Journal of Chinese Linguistics*. 10 , 189-206.