THE ACQUISITION OF MANDARIN AFFRICATES BY AMERICAN SECOND LANGUAGE LEARNERS*

Chunsheng Yang[1] and Alan C. L. Yu[2]
[1]The University of Connecticut
[2]The University of Chicago

ABSTRACT

The acquisition of new affricate contrasts in a second language (L2) often poses difficulties for L2 learners. This study examined the perception and production of six Mandarin affricates by beginning and intermediate American L2 learners using one production and two perceptual (i.e., identification and discrimination) tasks. The results of the perceptual tasks showed that the L2 learners performed as equally well as the native speakers in the discrimination task, but did not perform as well as the native speakers in the identification task, and that the difficulty in the identification task was not consonant-specific. The results of the production task showed that the affricates produced by the L2 learners differ from those produced by the native speakers along several acoustic dimensions (center of gravity, skewness, ampRatio, and frication duration). The study concludes by discussing the complex interaction between L2 speech production and perception, as well as the roles of the L1, L2, and L2 pedagogy in shaping instructed second language acquisition.

Key words: Affricate, Mandarin, assimilation, second language acquisition, salience

---

1. INTRODUCTION

When acquiring a phonological contrast in a second language (L2), learners may or may not encounter difficulties, depending on how the L2 contrast is assimilated into the first language (L1) (Best 1995; Best and Tyler 2007) or the degree of similarity between the L1 and L2 contrasts (Flege 1995; Flege et al.2003). Linguistic universals and markedness are also important factors affecting second language acquisition (Eckman 1977, 1991; Major 2001). Learning L2 sounds entails learning both the perception and the production of these sounds. Most studies on L2 phonological acquisition have shown that the perception of L2 sounds precedes L2 production presumably because only after the successful perception of L2 sounds can the relevant sensorimotor skills in L2 production be achieved (e.g., Flege, Schirru and MacKay 2003; Piske, MacKay and Flege 2001); listeners' performance in L2 perception tasks also tends to be better than their performance in L2 production tasks (Flege 1993; Flege, Bohn and Jang 1997; Flege, MacKay and Meador 1999; Neufeld 1988, among others). However, some studies, particularly with respect to the perception and production of the English /r/ and /l/ by Japanese speakers, have shown that the successful learning of L2 production may precede that of L2 perception (Goto 1971; Sheldon and Strange 1982).

This study contributes to the literature on L2 speech learning by examining the perception and production of Mandarin affricates by American English L2 learners. Assuming that the accurate production of L2 sounds requires L2 learners to accurately perceive the phonetic contrasts that may be present in a language, it is expected that the successful L2 production of Mandarin affricates should be preceded by accurate perception by the American English L2 learners. This study contributes to the empirical knowledge of L2 speech production and perception research. The literature on L2 phonetic acquisition has largely focused on European languages, such as English and French, and studies on the acquisition of L2 Mandarin phonetics with regard to perception and production are still in their infancy.

In recent years, the increase in the popularity of the study of Mandarin Chinese has been accompanied by a rise in the number of

studies on the acquisition of Mandarin Chinese. However, most of the studies on the acquisition of L2 Mandarin phonology have focused on the acquisition of lexical tones (Zhang 2013 and references therein), while studies on the acquisition of L2 Mandarin consonants and vowels are scant (Hao 201 focuses on the acquisition of L2 Chinese consonants, vowels and tones. Huang (2013) examines the acquisition of Chinese affricates by Spanish speakers). This study represents an important step to fill an empirical gap by examining the L2 acquisition of a particularly complex part of the Mandarin consonant inventory.

This study is organized as follows. We begin by offering some additional background regarding Mandarin affricates, followed by a review of previous findings on the L2 acquisition of these segments in Section 2. We then report the methodology (Section 3) and results (Section 4) of a series of perception and production experiments. In particular, we examine the perception of Mandarin affricates by American English-speaking L2 learners to see whether the L2 learners can correctly identify and discriminate between the Mandarin affricates when compared to the native speakers. We also examine L2 learners' production of these affricates by focusing on several defining acoustical parameters. Section 5 provides a discussion of the findings, with a focus on the relationship between L2 production and perception by drawing on the perceptual and production data.

2. BACKGROUND

2.1 Affricates in English and Mandarin Chinese

Affricates occur in over two-thirds of the world languages (van de Weijer and Hinskens 2004). However, the existence of affricates in many languages does not mean that they do not pose difficulty for L2 learners. Difficulties might arise due to the fact that the L1 and L2 may have different affricate systems. Affricates in one language may have different phonotactic constraints than those in another language (i.e., only occurring in some phonological contexts and not others). The acquisition of Mandarin affricates by English-speaking L2 learners is a

case in point. The affricates in Mandarin contrast in three places of articulation (Duanmu 2000): dental, retroflex (cf. post-alveolar in Lin2007:44), and palatal (alveolo-palatal in Lin 2007:47), while there is only one pair of post-alveolar affricates in English (Phonetically, [ts] and [dz] also exist in English, although only with a very restricted distribution, namely, as the consequence of morphology, e.g., ca[ts] and pa[dz]). Table 1 presents the inventory of affricates in English and Mandarin Chinese.

Table 1. Affricates in English and Mandarin Chinese

	Dental	Retroflex	Post-alveolar	Palatal
English	([ts] [dz])		[tʃ] [dʒ]	
Mandarin	z* [ts] c[tsh]	zh[tʂ] ch[tʂh]		j[tɕ] q[tɕh]

(*The *pinyin* Romanization for transcribing Chinese sounds will be used throughout this article to represent the Mandarin affricates)[1]

The precise nature of the retroflexes in Chinese is a matter of debate. Ladefoged and Maddieson (1996:150-3) claimed that Mandarin retroflexes are not apical but laminal in that the tongue blade, instead of tongue tip, is involved in the articulation. Lin (2007:46) disagreed with Ladegfoged and Maddieson and argued that, if it were the case, the English [ʃ] and the Chinese [ʂ] would be produced in the same way because both are laminal post-alveolar. However, it is worth mentioning that the difference in the descriptions of the retroflexes of Lin and of Ladefoged and Maddison might stem from a difference in the variety of Mandarin under discussion. What Ladefoged and Maddieson referred to is the pronunciation of Beijing Mandarin, while Lin's description might

---

[1] One reviewer pointed out that the difference between Chinese retroflexes and palatals are allophonic, not phonemic, and that the phonological status of these two pairs of affricates may bear upon the results of this study. While we agree with the reviewer on such difference, we do not feel that it will affect the results of the study, because both native speakers and L2 learners are able to differentiate retroflexes from palatals, due to the difference in *pinyin Romanization* and the explicit teaching. We do agree that the allophonetic status of Chinese retroflexes and palatals may make it more difficult for L2 learners to acquire.

reflect the retroflexes produced by speakers of Mandarin whose first language is another dialect of Chinese, but not Beijing Mandarin. Another important difference between Mandarin and English affricates has to do with their phonotactic distribution. Mandarin affricates only occur at the syllable-initial position, while English affricates occur at both syllable-initial and syllable-final positions. What makes the acquisition of Mandarin affricates by L2 English speakers more complicated might be that, as noted above, while dental affricates do not occur contrastively in English, they do occur at the word-final position as a stop-fricative sequence, such as in *ts* in *cats* and ds in *beds*. According to Johnson (2002), the main acoustical difference between an affricate and a stop-fricative sequence is that the amplitude of the frication noise rises quickly to full amplitude in the case of affricates, while it rises more slowly in the case of fricatives. Whether such nuanced acoustical differences can be perceived or produced by L2 learner remains an open question, though. One last difference between Mandarin and English affricates lies in the contrast for voicing and aspiration. English affricates contrast for voicing, although the voiceless affricates in English can be aspirated or unaspirated, depending on the phonological context, while Chinese affricates only contrast for aspiration. As a result, the acquisition of Mandarin affricates is expected to pose difficulty for American English-speaking L2 learners of Mandarin since the allophonic differences in English (L1) may be mapped onto the contrastive differences in Mandarin (L2).

2.2 Studies on the Acquisition of Mandarin Affricates by L2 Learners

Earlier studies on the perception and production of Mandarin affricates by L2 learners have focused on L2 learners whose native languages are other than English. L2 learners whose native languages use contrast for aspiration have been found to be able to employ the aspiration feature in their L1s to assimilate the Chinese affricates into their native sound systems (Mei 2011). Lai (2009) reported that some learners had great difficulties with the perception of the dental-retroflex contrast. Indonesian learners, whose affricates only contrast for voicing,

but not aspiration, tended to assimilate the aspiration contrast into the voicing contrast (Lin 2006; Lin and Wang 2005; Si 2011).

Liu and Jongman (2012) investigated American learners' acquisition of Mandarin affricates, by examining the temporal and spectral properties of the Chinese dental affricates ([ts] and [tsʰ]) produced by American English L2 learners of Chinese and found that the novice learners (who had learned Chinese for less than six months) only acquired the duration contrast between this pair of affricates, but did not acquire the difference in center of gravity (CoG) between them. The advanced learners (who had learned Chinese for about 2.5 years) had acquired both the duration and the CoG contrasts. Liu and Jongman (2012) argued that their findings suggested a later acquisition of the spectral contrast for the L2 learners, relative to the acquisition of the duration differences.

The present study builds on these earlier findings and looks at perception and production combined. We examine the perception and production of Mandarin affricates by native speakers of Mandarin as well as by beginning- and intermediate-level L2 Mandarin learners. Unlike the earlier studies, which have generally focused on a subset of the Mandarin affricates, we examine the L2 learners' perception and production of all of the six affricates (i.e., [ts], [tsʰ], [tʂ], [tʂʰ], [tɕ], [tɕʰ]) in Mandarin.

2.3 Speech Learning Model (SLM) and Perceptual Assimilation Model (PAM)

Various theoretical models have been proposed for L2 speech learning, such as the speech learning model (SLM) and the perceptual assimilation model-L2 (PAM-L2). According to the SLM (Flege 1995), an L2 sound may be identical, similar, or new, when compared to L1 sounds. The identical sounds are easy to both perceive and produce due to L1 positive transfer (Weinreich 1953). Similar L2 sounds, however, may be difficult to perceive and produce because the similar L2 sound may be equivalently classified as an L1 sound without more detailed refining. By contrast, L2 learners may establish new sound categories for sounds which are new to them. However, it should be noted that, while

L2 learners may establish such sound categories for new sounds, their production of such sounds may not be as good as their production of the identical sounds, in that the established L2 categories may not approximate to the L1 counterparts.

The perceptual assimilation model (PAM) for experienced L2 learners (PAM-L2) (Best 1995; Best and Tyler 2007) is another important theoretical model on non-native speech perception. Depending on whether a learner perceptually assimilates one L2 phone to his/her phonological entity, PAM-L2 predicts five possible scenarios of how L1 and L2 sounds relate at the phonological level: two-category assimilation (i.e., both L2 phones are assimilated to different L1 phonological categories; good discrimination), single-category assimilation (i.e., both L2 phones are perceptually assimilated to the same L1 phonological category, but as equally good or poor instances of that category; poor discrimination), category-goodness assimilation (i.e., both L2 phones are assimilated to the same L1 phonological category, but one is perceived as more deviant than the other. Discrimination of the two L2 sounds will be very good), uncategorized-categorized assimilation (i.e., one L2 phone is assimilated to one L1 category, and the other cannot be perceptually assimilated to any L1 phonological category; good discrimination), and uncategorized assimilation (i.e., neither of the L2 phones can be perceptually assimilated to the L1 phonological category; easy to discriminate).

2.4 Research Questions, Hypotheses and Predictions

This study aims to examine the following three research questions:
1) Do American L2 learners at the beginning and intermediate levels acquire the three pairs of Mandarin affricates in the same way? If not, what is the difference in production in terms of different acoustic parameters and in perception across the groups?
2) How can we explain the difference in the production and perception of different Mandarin affricates by different groups?
3) What new light can the L2 affricate production and perception data in this study shed upon the relationship between L2 speech production and perception?

The following hypotheses and predictions are proposed based on previous studies and the SLM and PAM-L2.

For the perception of Chinese affricates:

Hypothesis 1a: Since Chinese affricates contrast for aspiration whereas English ones do not and previous studies (Lin 2006; Lin and Wang 2005; Si 2011) have shown that American English L2 learners of Chinese tend to assimilate the aspiration contrast into the voicing contrast, it is predicted that L2 learners should not face difficulty in discriminating the two affricates with the same place of articulation.

Hypothesis 1b: English only has two post-alveolar affricates, as compared to the three pairs of affricates in Chinese, and, thus, according to the SLM and PAM-L2, Chinese retroflexes and palatals may be perceptually assimilated to English post-alveolar affricates. Therefore, it is predicted that the discrimination (one category assimilation in PAM-L2) and the identification of the retroflex-palatal pairs of affricates may be difficult for L2 learners.

For the production of Chinese affricates:

Hypothesis 2: If L2 learners assimilate Chinese retroflexes and palatals to English post-alveolar affricates for the production of Chinese affricates due to their similarity, it is predicted that there will be no or little difference in the different acoustical parameters between the L2 retroflexes and palatals.

3. METHODOLOGY

3.1 Subjects

Three cohorts were recruited. Eight beginning level (who had learned Chinese for three months at the time of recording) and eight intermediate level (who had learned Chinese for 1.5-2.5 years) L2 learners of Mandarin Chinese were recruited as the test groups, while eight Beijing Mandarin speakers were recruited as the control. The Beijing Mandarin speakers were born and grew up in Beijing before coming to study in the US. All of the participants were students at a university in the New

England area of the USA. Table 2 provides more information of the participants.

Table 2. Demographics of the three groups of participants

Group	Mean age (SD)	Male	Female	Duration of Chinese learning (SD)
Beginning	19.3 (1.8)	5	3	0.25 (0)
Intermediate	20.4 (2.5)	2	6	2.3 (0.2)
Native	25.6 (3.2)	4	4	n/a

3. 2 Stimuli

The speech stimuli used in this study are zi [tsɹ], ci [tsʰɹ], zhi [tʂɻ], chi [tʂʰ ɻ], ji [tɕi], and qi [tɕʰi]. The three vowels, the apical [ɹ], the retroflex [ɻ], and the high front vowel [i], are in complementary distribution in Mandarin Chinese. More specifically, [ɹ] occurs after the dental affricates and fricatives, [ɻ] occurs after the retroflex affricates, fricatives, and approximant ([ɻ]), and [i] occurs after the palatal affricates and elsewhere. These three vowels share the same *pinyin* Romanization *i* in Mandarin, although L2 learners are usually reminded that the pinyin *i* has different variants depending on the phonological contexts. To avoid any potential effects of tones on affricate perception and production, all of the target syllables carry (carried a high level tone (i.e., Tone 1 in Mandarin Chinese). One male Beijing Mandarin speaker in her late thirties recorded each of the target syllables and the six fillers ([ni, mi, ti, di, li, pi]) three times. In the discrimination and identification tasks, only stimuli with a similar length of duration (~500 ms) were used. The intensity of each of the stimuli was normalized. The F0 was resynthesized to make sure that the test syllables had the same pitch height and contour.

3.3 Design and Procedure

The letters of participant recruitment were posted on campus bulletin boards and in several student mail lists of the university. After participants contacted the researchers, an appointment was made. At the meeting, the participants were first briefed about the research. If they agreed to continue to participate, they signed the written consent form. Then they filled in the demographic form about their personal information. The two perceptual tasks and one production task then followed. The perceptual tasks preceded the production task. The two perceptual tasks were run on E-prime, while the stimuli for the production task were presented via a web browser using a customized Java script. The whole session lasted less than one hour and each participant was paid $10 for their participation.

In one perceptual task (i.e., the AX discrimination task), participants were presented with pairs of CV syllables where the two syllables were separated by an inter-stimulus interval of 100 ms. The participants were instructed to indicate whether the two syllables were the same different using buttons labeled SAME and DIFFERENT on the computer screen. The button positions were counter-balanced. 120 trials (30 target discrimination pairs x 2 within-trial order of the CV pairs + 10 repetitions of the 6 identical pairs) were randomly presented within a single block and the trial block was repeated four times.

In the other perceptual task (i.e., the identification task), the participants listened to the target syllables in isolation and were asked to identify each syllable as one of the six possible responses in *pinyin* Romanization. Each target syllable was repeated ten times, hence 60 trials. The participants were given three seconds to respond to each stimulus. Before the perceptual tasks were begun, the participants were informed that they needed to stay alert and respond as quickly as possible in both perceptual tasks. Both perceptual tasks were implemented in E-Prime and the participants' responses and reaction time were automatically recorded in the program.

The production task was administered immediately after the two perceptual tasks. The target and filler syllables were embedded in a carrier sentence, "请说___字 [tɕʰiŋ ʂuɔ __ tsɻ]" "Please say the __

character". Both Chinese characters and *pinyin* Romanization were provided in case that some L2 learners might not know the Chinese character. The target syllables were 資 zi [tsɿ], 疵 ci [tsʰɿ], 知 zhi [tʂʅ], 吃 chi [tʂʰʅ], 雞 ji [tɕi],and 七 qi [tɕʰi] and the fillers were ni [ni], mi[mi], ti [tʰi], di [ti], li[li], and pi [pʰi]. Each target syllable was repeated six times in the production task. A custom-made Java script presented each target sentence one at a time in random order via a web browser. Six repetitions were elicited in six blocks, each block contained one full randomized set of the target sentences. The recording was recorded with Audacity, a free audio recording software, directly onto a computer with a head-mounted microphone in a quiet room.

3.4 Data Measurement

This study involves both perceptual and production tasks. For the two perceptual tasks, reaction time and response accuracy were recorded. In the production task, the first four spectral moments, namely the spectral mean center of gravity (CoG henceforth), standard deviation, skewness, and kurtosis of the affricates produced by the Chinese native speakers and the L2 learners were measured. In addition, the duration of the frication noise and the amplitude ratio (ampRatio henceforth) were measured as well. The ampRatio, defined as the difference in dB between the F2 amplitude and the amplitude of the most prominent peak above the F2 region, is used to assess the degree of "palatalization" (i.e., the difference in tongue posture) (Li, Edwards and Beckman 2007).

For the two perceptual tasks, the reaction time and response accuracy were recorded. For the production task, the frication interval of the affricates were labeled manually and a custom-made Praat script automatically extracted the spectral measurements, namely the CoG, standard deviation, skewness, and kurtosis, using a 40 ms Hamming window with the preemphasis at 80 hz, centered at the midpoint of the frication portion of the affricate. The same script also measured the duration of the frication noise, and the amplitude ratio. The ampRatio was calculated to assess the degree of "palatalization". We determined the average amplitude within the F2 region, which was estimated by taking the F2 at the onset of the following vowel and defining a 1000 Hz

band around that peak. The average amplitude within the F2 region was then estimated by subtracting the average amplitude within the F2 region from the average amplitude of a 1000 Hz band center on the highest amplitude peak above the F2 region.

3.5 Statistical Analysis

The continuous variables (e.g., reaction time, CoG, standard deviation, kurtosis, skewness, ampRatio, and frication duration) were modeled using linear mixed-effects regression fitted in R, using the lmer () function from the lme4 package (Bates et al. 2011), and the p-values were obtained using normal approximation (Mirman 2014). The binary variables (e.g., identification and discrimination accuracy) were modeled using a logistic mixed-effects regression. The Wald's Z test, which describes how distant a coefficient estimate is from zero in terms of its standard error, was used to test the significance of the estimates of the model. To reduce the multicollinearity between the predictors, the continuous variables were centered and scaled, while the categorical variables were sum-coded unless mentioned otherwise.

4. RESULTS

In what follows, we report the results of the three experiments. We begin by reporting the results of the identification and discrimination experiments, followed by the acoustic results of the production task.

4.1 Results of the Two Perceptual Tasks

Results of the identification task

The participant's reaction time for the correct responses was (natural) log-transformed before being submitted to regression modeling. The following within-subject predictors were tested: trial order (1-38), CONSONANTTYPE (dental vs. palatal vs. retroflex), and ASPIRATION (unaspirated vs. aspirated). GROUP was also included as a between-subject factor. CONSONANTTYPE and GROUP were Helmert-coded. For

GROUP, contrast 1 compares the performance of the native speakers to that of the learners, while contrast 2 compares the performance of the two learner groups. For CONSONANTTYPE, contrast 1 compares the dentals to the other two places of articulations, while contrast 2 compares between retroflexes and palatals.

In both the reaction time and identification accuracy models, all two-way interactions between CONSONANTTYPE, ASPIRATION, and GROUP were included. The addition of a three-way interaction between these three predictors did not significantly improve model-likelihood and was therefore not included in the final regression models. The models included by-subject random intercepts, to allow for subject-specific variation in the specific acoustic measure as well as by-subject random slopes for TRIAL CONSONANTTYPE, and ASPIRATION to allow for by-subject variability in the interaction between these factors. Random slopes for the interaction of CONSONANTTYPE, and ASPIRATION were tested but that model did not converge and was thus not included in the final analysis.

Table 3 summarizes the parameter estimates for each of the fixed effects in the regression models for identification reaction time and accuracy respectively, while Figure 1 illustrates the reaction times (left) and the accuracy rates (right) by consonant type and speaker group. To begin with, the native speakers responded faster than the learners (β=-0.18, z=-3.95, $p < 0.001$). There are also significant differences in the response term, depending on the place of articulation of the consonant. The responses in the case of the dentals were slower than those in the cases of the other two places of articulation (β=0.04, z=2.26, $p < 0.05$), while the responses to the retroflexed affricates are the fastest (β=-0.07, z=-3.52, $p < 0.001$). The place-dependent difference in response time differs across groups. While the learners are generally slow in their responses, the native speakers are particularly quick in their responses to the retroflexed affricates (β=-0.10, z=-2.45, $p = 0.01$). Between the two learner groups, the beginners are particularly slow in their responses to the dental affricates (β=0.13, z=2.92, $p < 0.001$).

Table 3. Main and interaction effects of all predictors in the reaction time and accuracy models for the identification experiment

Predictor	Reaction Time			Accuracy		
	Coef	SE		Coef	SE	
Intercept	7.301	0.022	***	4.080	0.398	***
Trial	-0.009	0.008		0.317	0.175	
ConsType1	0.041	0.018	*	1.246	0.653	
ConsType2	-0.069	0.019	***	1.082	0.651	
Aspiration	0.009	0.008		0.235	0.296	
Group1	-0.179	0.045	***	3.356	0.998	***
Group2	0.016	0.053		-1.252	0.696	
ConsType1:Aspiration	-0.001	0.014		0.210	0.354	
ConsType2:Aspiration	-0.025	0.017		1.300	0.272	***
ConsType1:Group1	0.014	0.036		-1.863	1.573	
ConsType2:Group1	-0.096	0.039	*	2.587	1.750	
ConsType1:Group2	0.126	0.043	***	-0.519	1.287	
ConsType2:Group2	0.079	0.049		-0.210	0.957	
Aspiration:Group1	0.010	0.017		1.395	0.766	
Aspiration:Group2	-0.034	0.020		0.463	0.433	

($* p < 0.05$. $** p < 0.01$, $*** p < 0.001$; GROUP1 compares the native speakers to the learners, while GROUP2 compares between the two learner groups. CONSONANTTYPE1 compares the dentals to the other two places of articulation, while CONSONANTTYPE2 compares between retroflexes and palatals)
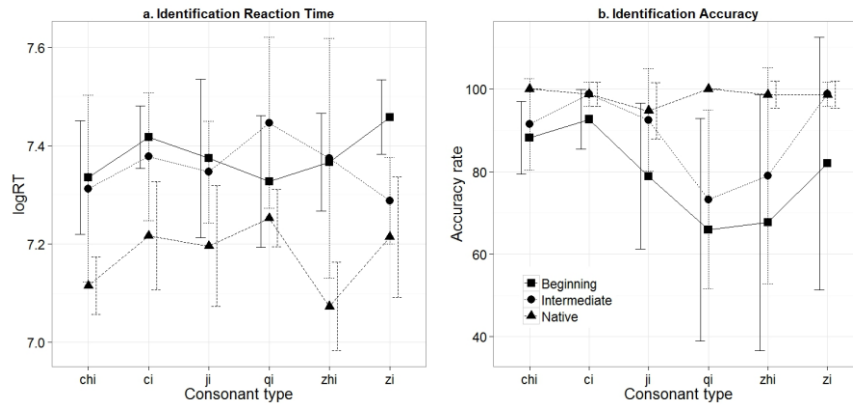
Figure 1. Log-transformed reaction time (a) and accuracy in the identification task across consonant type and speaker group (b). The error bars indicate 95% confidence intervals.

In terms of identification accuracy, the native speakers also performed significantly better than the learners (β=3.36, z=3.36, $p <$ 0.001). There is also an interaction between ConsonantType and Aspiration (β=3.36, z=4.78, $p <$ 0.001). In particular, identification accuracy varies depending on the aspiration of the affricate in addition to whether the consonant is palatal or retroflex. As shown in Figure 2, the accuracy of both the aspirated palatal q /tɕʰ/ and the unaspirated retroflex zh /tʂ/ were markedly lower when compared to those of their opposite aspiration counterparts ((i.e., j /tɕ/ and tʂ /tʂʰ/, respectively).
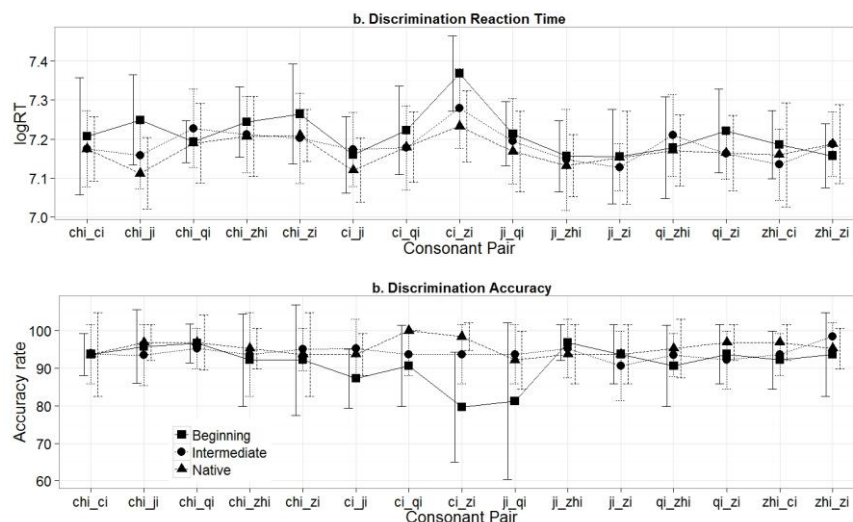
Figure 2. Log-transformed reaction time (a) and accuracy rates (b) in the discrimination task across consonant type and speaker group.

Results of the AX discrimination task

The accuracy of "different" pairs was analyzed using a logistic mixed-effects model. The discrimination pairs were analyzed in terms of the place of articulation and aspiration of the affricate pairs. PAIRTYPE(PLACE) has six levels: dental-dental, dental-palatal, dental-retroflex, palatal-palatal, retroflex-palatal, and retroflex-retroflex. PAIRTYPE(ASPIRATION) has two levels. The two affricates in a discrimination pair either share the same aspiration or have different aspiration. Figure 2 summarizes the discrimination response time and accuracy results. Besides pair type, trial order (1-240) was also entered as a within-subject predictor. In the end, no significant difference in discrimination accuracy across discrimination pairs, GROUP, nor any interaction between these two predictors was found.

In terms of response time, the only significant predictor of logRT is PAIRTYPE(PLACE). Specifically, the response time is significantly longer than average when both affricates are dental (i.e., ci~zi pairs; β=0.08, z=6.18, $p < 0.001$). Response time is significantly shorter than average when the affricate pairs are dental-palatal (β=-0.04, z=-4.84, $p < 0.001$),

dental-retroflex (β=-0.02, z=-2.13, $p < 0.05$), and retroflex-palatal (β=-0.03, z=-4.02, $p < 0.001$). There was no significant GROUP nor PAIRTYPE(APIRATION) difference. The addition of the interaction of GROUP with the two PairType predictors did not significantly improve the model likelihood

Taken together, the results of the AX discrimination task suggest that participants are similar in their abilities to discriminate between Mandarin affricates, regardless of their level of proficiency. In terms of consonant type, listeners in general found it more difficult to distinguish the dental affricates.

4.2 Results of the Production Task

The CoG, standard deviation, kurtosis, skewness, ampRatio and frication duration were modeled separately using linear mixed-effects regression fitted in R. Because of the heavily skewed distribution of the kurtosis parameter, the kurtosis measure was log-transformed (i.e., log(x + 2) where x is the actual kurtosis value) for further statistical analysis. We also added two to all the kurtosis values to ensure that they were positive prior to log-transformation. All (acoustic measures were tested for the following within-subject effects: trial order (1-38), CONSONANTTYPE (dental vs. palatal vs. retroflex), and ASPIRATION (unaspirated vs. aspirated). GROUP was also included as a between-subject factor. CONSONANTTYPE and GROUP were Helmert-coded. For GROUP, contrast 1 compares the performance of the native speakers to the mean performance of the two learner groups, while contrast 2 compares between the two learner groups. For CONSONANTTYPE, contrast 1 compares the dentals to the other two places of articulations, while contrast 2 compares between the retroflexes and the palatals.

The models also included by-subject random intercepts, to allow for subject-specific variation in the specific acoustic measure as well as by-subject random slopes for TRIAL, CONSONANTTYPE, ASPIRATION, and the interaction of the latter two to allow for by-subject variability in the interaction between these factors.

The residuals of the initial fit of each model were examined and were found to deviate strongly from normality. As a result, the residuals which

were more than 2.5 standard deviations from the mean were trimmed, which amounted to no more than 3% of the data, and the models were refitted to the trimmed data set. The new model had a residual distribution much closer to normality, and it is the refitted models that are reported below.

Table 4 provides a summary of the main and interaction effects of all the predictors for each acoustic measure. Only significant predictors at the $p < 0.05$ level are reported. The average spectral mean (CoG) for the frication noise of the affricates is approximately 6504 Hz. There is a main effect of GROUP where the native speakers have a higher spectral mean than the L2 learners (β=955.51, $t = 2.61$, $p < 0.01$), but the difference between the learner groups is not significant. There is no main effect of ASPIRATION, but there is a significant interaction between GROUP and ASPIRATION (β=-152.69, $t = -3.19$, $p < 0.01$). Native speakers have higher CoG for the unaspirated sound than the aspirated ones, while the L2 learners show the opposite pattern. There are main effects of CONSONANTTYPE. Specifically, the dental affricates have the highest CoG relative to the other two places of articulation (β=3285.40, $t = 15.85$, $p < 0.001$). Between the retroflex and palatal affricates, the retroflex has the lower spectral mean (β=-1448.33, $t = -7.51$, $p < 0.001$). The CONSONANTTYPE effect differs by GROUP, however. As shown in the top left panel of Figure 3, the native speakers produced clear differences in the CoG between the palatals and retroflexes (β=-2625.16, $t = -6.58$, $p < 0.001$), but the L2 learners did not show such marked differences. The difference in the CoG across CONSONANTTYPE is mediated by aspiration; the spectral mean of the retroflex affricates is higher when the affricates are aspirated (β=93.81, $t =2.25$, $p < 0.05$), while the CoG of the palatals shows the reverse pattern with respect to aspiration. Finally, there are significant three-way interactions as well. As shown in Figure 3, while the native Mandarin speakers show higher CoG values for unaspirated palatals, the learners exhibit no clear difference between aspirated and unaspirated palatals (β=565.4, $t =6.57$, $p < 0.001$). On the other hand, the beginners exhibit a lower CoG among the unaspirated dentals compared to their aspirated counterparts while the intermediate learners show no such difference (β=-308.6, $t =-3.10$, $p < 0.01$).

Table 4. Main and interaction effects of all (Only the coefficients of the significant effects are given, * p < 0.05. ** p < 0.01)

Predictor	CoG	Std	Skewness	Kurtosis
Intercept	6504 **	1600 **	0.55 **	1.09 **
Trial				
Aspiration				-0.04 *
ConsType1: Dental vs. the rest	3285 **		-1.44 **	-0.45 **
ConsType2: Retroflex vs. Palatal	-1488 **	108 *		
Group1: Native vs. Learners	956 *		-0.70 **	
Group2: Intermediates vs. Beginners				
Aspiration: ConsType1				-0.10 **
Aspiration: ConsType2	94 *			
Group1: Aspiration	-153 **		0.19 **	
Group2: Aspiration				
Group1: ConsType1				0.73 *
Group1: ConsType2	-2625 **		1.54 **	0.62 *
Group2: ConsType1				
Group2: ConsType2				
Group1: Aspiration: ConsType1				
Group1: Aspiration: ConsType2	565 **		-0.33 **	
Group2: Aspiration: ConsType1	-309 **			
Group2: Aspiration: ConsType2				

Table 4. Main and interaction effects of all (Only the coefficients of the significant effects are given, * p < 0.05. ** p < 0.01)

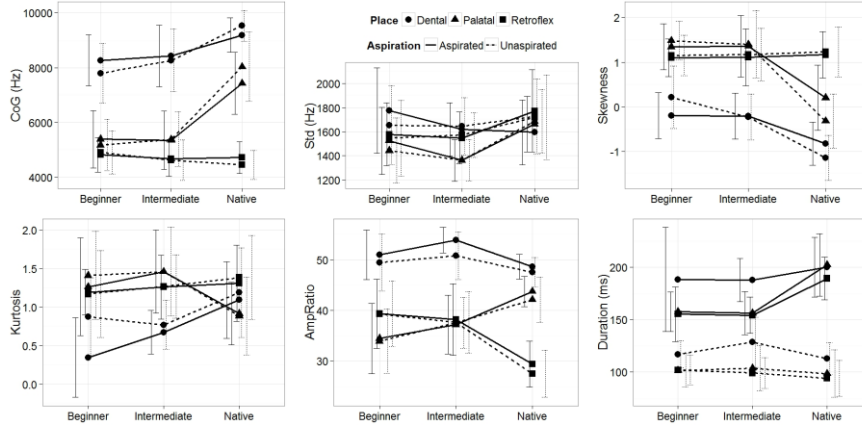|	ampRatio		Duration	
Predictor	Coef		Coef	
Intercept	41.14	**	141.48	**
Trial			-2.94	*
Aspiration	0.69	**	35.26	**
ConsType1: Dental vs. the rest	13.38	**	21.64	**
ConsType2: Palatal vs. Retroflex	-3.08	*		
Group1: Native vs. Learners				
Group2: Intermediates vs. Beginners				
Aspiration:ConsType1				
Aspiration:ConsType2				
Group1:Aspiration			18.46	**
Group2:Aspiration				
Group1:ConsType1			-17.91	*
Group1:ConsType2	-17.97	**		
Group2:ConsType1				
Group2:ConsType2				
Group1:Aspiration:ConsType1				
Group1:Aspiration:ConsType2				
Group2:Aspiration:ConsType1				
Group2:Aspiration:ConsType2				

Figure 3. Center of gravity (CoG), standard deviation (Std), skewness, kurtosis, ampRaio, and frication duration for each Mandarin affricate place of articulation and aspiration type across speaker groups.

In terms of the standard deviation, the only significant effect is in the place of articulation (β=108.32, t = 2.24, $p < 0.05$). In particular, the spectral energy is more diffused with respect to the retroflexes than the palatals.

Concerning skewness, which characterizes the slant of the spectral energy distribution, the dentals are significantly more negatively skewed with a high concentration of spectral energy in the higher frequencies) than the retroflexed and palatal affricates (β =-1.44, t = -9.14, $p < 0.001$). The native speakers also exhibit a more negative skew than the learners (β =-0.70, t = -3.81, $p < 0.001$). In particular, relative to the native speakers, the L2 learners generally *shows* a more positive skew for the unaspirated affricates than for the aspirated ones (β =0.19, t = 4.34, $p < 0.001$), and they show less positive skew for the retroflexes than the palatals (β =1.54, t = 4.86, $p < 0.001$). As shown in the top right panel of Figure 3, the palatal affricates have much less positive skewness than retroflex affricates among the native speakers and this negative difference between the palatals and the retroflexes among the native speakers is stronger when the affricates are unaspirated (β =-0.33, t = -3.61, $p < 0.01$).

In terms of kurtosis, the frication noise of the unaspirated affricates has a peakier distribution than their aspirated counterparts (β =-0.04, t = -2.53, $p < 0.05$). The dentals also show significantly less peaky spectral profile than the retroflexes and palatals (β =-0.45, t = -3.08, $p < 0.01$). There is a significant interaction between ASPIRATION and CONSONANTTYPE (β =-0.10, t = -2.89, $p < 0.01$). In particular, the aspirated dentals show particularly smaller kurtosis values compared to the unaspirated dentals. The interaction between GROUP and CONSONANTTYPE is also significant. As illustrated in the bottom left panel of Figure 3, the L2 learners show generally higher kurtosis values for the retroflexes and palatals compared to the dentals, while native speakers exhibit (a more complicated picture. In particular, native speakers show a higher kurtosis value for the dentals compared to the nondentals (β =0.73, t = 2.37, $p < 0.05$), but crucially, for the native speakers, the kurtosis of the palatal affricates is significantly lower than that of the retroflexes (β=0.62, t = 2.57, $p < 0.05$).

There are also significant differences in terms of the ampRatio. To begin with, there is a significant ampRatio difference between the aspirated and unaspirated affricates (β =0.69, t = 3.26, $p = 0.001$); the unaspirated affricates tend to have lower ampRatios. The place of articulation also affects the ampRatios in significant ways; the dentals have higher ampRatios than the nondentals (β =13.38, t = 9.68, $p < 0.001$), while the palatals have the lowest ampRatios (β =-3.08, t = -2.33, $p < 0.05$). The ampRatio difference between the retroflexes and palatals is mediated by GROUP (β =-17.97 t = -6.47, $p < 0.001$); the retroflex affricates have much lower ampRatios than the palatals among the native speakers, but the ampRatio difference is much smaller for the learners.

The frication duration, as expected, differs, depending on the aspiration type (β = 35.26, t = 17.44, $p < 0.001$); the frication duration is approximately 35 ms longer in the aspirated affricates relative to the unaspirated affricates. The frication duration is significantly longer among the dental affricates compared to the rest (β =21.64, t = 5.91, $p < 0.001$). The aspirated affricates have longer frication duration for the native speakers than for the learners (β =18.46, t = 4.34, $p < 0.001$). While the learners exhibit a longer frication duration among the aspirated dentals, the frication duration does not differ significantly

across the aspirated affricates of the different place of articulation for the native Mandarin speakers (β =-17.91, t = -2.31, $p < 0.05$). Finally, there is a main effect of Trial, suggesting that frication duration decreased as the experiment progressed (β =-2.94, t = -2.39, $p < 0.05$).

5. SUMMARY AND DISCUSSIONS

In this study, two perceptual and one production tasks were designed to examine the acquisition of three pairs of Mandarin affricates by American English L2 learners. The discrimination task showed that all participants could discriminate the affricates relatively well, even if there exist differences in the response time depending on the place of articulation of the affricates (for example, the learners found dentals more difficult to distinguish in general). The results of the identification task showed that the native speakers overall did better overall than the two learner groups in correctly identifying the affricates. There was no significant difference in accuracy across the two learner groups nor was there significant effect of consonant type across groups, even though beginners were generally slower in their responses to dental affricates and native speakers were particularly quick in identifying retroflex affricates. Therefore, Hypothesis 1a was confirmed. The learners' identification accuracy varied depending on the aspiration of the affricate and the place of articulation. For example, the aspirated palatal q /tɕʰ/ and the unaspirated retroflex zh /tʂ/ have lower identification accuracy than their opposite aspiration counterparts (i.e., j /tɕ/ and ch /tsʰ/), which is difficult to interpret in that the identification difficulty is not directly linked to aspiration or the place of articulation. Meanwhile, the results of the discrimination tasks did not show significant difference across affricate pairs. Thus, these findings do not support predictions of Hypothesis 1b. The difference in the results of the discrimination and identification tasks may be due to the nature of the two tasks. The discrimination task mainly taps into the phonetic difference, namely, as long as the listeners heard any phonetic/acoustic difference, they would be able to discriminate them, which could explain the non-difference in discrimination across groups. However, the identification task taps into

the phonological knowledge of listeners and, as a result, can pose difficulty for learners to varying degrees, which could explain the better identification by the native speakers than the two learner groups.

The analysis of the production tasks shows that the native speakers could be differentiated from the L2 learners in the case of all of the acoustical parameters although the two learner groups could not be further differentiated from each other in general. Thus, Hypothesis 2 was mostly confirmed: both learner groups tended to assimilate Chinese retroflexes and palatals to English post-alveolar affricates, due to their phonetic similarity, as shown in the similarity of the CoG, skewness, kurtosis, and ampRatios of the palatals and retroflexes for the two learner groups (but not the frication duration). It seems that the learners have not completely acquired the Mandarin affricates.

While the performances of the two learner groups did not differ drastically, the differences in some acoustical parameters showed that there is learning effect between these two groups. For example, the beginners exhibited a lower CoG among the unaspirated dentals as compared to their aspirated counterparts, while no such difference was seen in the case of the intermediate learners, and the native speakers have higher CoG for the unaspirated affricates than the aspirated ones. In this sense, the intermediate learners approximated the performance of the native speakers more than the beginners, suggesting a learning effect, which also supports Liu and Jongman's (2012) findings. A similar pattern can be observed in the skewness for the dentals across the three groups. Also, the beginners produced the lowest ampRatio for the palatals, while the native speakers produced the lowest ampRatio for the retroflexes and the intermediate learners produced similar ampRatio for both the retroflexes and the palatals. These results of the ampRatio patterns across groups also suggest a learning effect. Specifically, the intermediate learners have lower ampRatio for the retroflexes and higher ampRatio for the palatals as compared to the beginners.

The across-group comparison of the affricate production revealed some mismatch between the three pairs of Mandarin affricates as well. For example, the frication noise of the L2 learners' affricates show a less positive skew for the retroflexes than the palatals, while there is a much less positive skew for the native speakers in the case of the palatals,

rather than the retroflexes. Further, while the L2 learners show a more negative skew for the aspirated affricates, especially the dentals, relative to the unaspirated ones, the native speakers have a more negative skew for the unaspirated affricates, both dentals and palatals, relative to the aspirated ones. The aspirated dentals of the learners show a longer frication duration than the non-dentals, while the results for the frication duration itself do not show any distinction between the aspirated affricates of different places of articulation for the native speakers. As shown in Figure 3, the frication duration for the aspirated dentals does not differ across groups, but the learners' non-dental aspirated affricates have shorter frication duration compared to the native speakers. That is, the learners had not acquired the aspiration feature in some aspirated affricates, such as retroflexes and palatals; however, they did acquire the aspiration in the dental affricates. Such mismatches show that different affricates pose different learning difficulties for L2 learners.

The difference in the frication duration between the aspirated dentals and the aspirated retroflexes and palatals for the learners may be related to Chinese pedagogy or the issue of salience in SLA (i.e., the availability of input; Gass, Behney and Plonsky 2013:148). American L2 learners usually have difficulties with the pair of dental affricates at the initial stage, specifically with respect to aspiration and voicing. Therefore, L2 Chinese instructors and Chinese textbooks usually associate this pair of affricates with the English equivalents ([ts] as in *cats* and [dz] "ds" as in *beds*), due to their phonetic similarity to Chinese dentals, and point out that the difference in the Chinese pair lies in aspiration, but not in voicing as in English. It seems that this particular way of teaching has served to facilitate the American L2 learners' acquisition of the aspiration in the aspirated dentals. Si (2011) argued that English L2 learners did not assimilate the Mandarin dental affricates to the English [ts^h] and [ts] because [ts^h] and [ts] are not independent phonemes in English. The findings in this study run counter to Si's argument in that the learners seemed to have assimilated the Chinese dental affricates to the English [ts^h] and [ts], due to L2 learning experience, and, as a result, acquired the aspiration feature in the dental affricates.

The findings of this study also show that the distinction between palatals and retroflexes is more difficult for the learners. It was argued

that the similarity of the Chinese retroflexes and palatals led the L2 learners to assimilate both affricates to the English post-alveolar affricates, hence their difficulty in differentiating these affricates. This result supports both the prediction of the SLM, namely, similar sounds pose more difficulty for L2 learners, and the prediction of the PAM-L2, namely, when two L2 sounds are assimilated to one L1 sound, the distinction of the L2 pair is very difficult.

Similar to Liu and Jongman (2012), we observed effects of learning in both production and perceptual tests, but with only marginal differences. This marginal effect of learning might be attributed to the fact that the learners did not vary considerably in terms of the duration of learning, as the beginners had studied Chinese for one semester, and the intermediate learners had studied Chinese for three or four semesters. Further, the small sample size may be another factor leading to the lack of a significant learning effect.

With respect to affricate perception, the findings of this study did not replicate the dental-retroflex merger reported in Lai (2009). This difference may be attributed to the variety of Mandarin learned by the L2 learners. Dental and retroflex affricates are not clearly distinguished by Taiwan Mandarin speakers. Since the L2 subjects in Lai (2009) were recruited in Taiwan, if the production of the dental and retroflex affricates are almost identical among the native speakers of Mandarin Chinese in Taiwan, it is not surprising that no such distinction is made by the L2 learners in Taiwan either. By contrast**,** the dental-retroflex affricate pairs are emphasized in almost all Chinese textbooks for Chinese as a Foreign Language (CFL), such as in *Integrated Chines*e (Liu, Yao, Ge, Chen, Bi, Wang and Shi 2008), *New Practical Chinese Reader* (Liu 2003)*, Chinese Link* (Wu, Yu, Zhang and Tian 2010)*,* and so on. The emphasis on the distinction of these affricates may explain why no dental-retroflex merger was found in this study.

The findings of this study show that L2 production and perception have a very complicated relationship. Although correct production entails an ability to identify the relevant acoustical cues, correct perception does not always lead to the correct production of the L2 sounds. The ways that L2 sounds are assimilated to L1 sounds influence the production of L2 sounds as well. According to the PAM-L2, if two

L2 sounds are assimilated to one L1 category, the ability to identify the two L2 sounds separately will be very difficult, as will the production of the sounds, because the L2 learners may employ similar acoustical cues to produce the two L2 sounds. What complicates L2 acquisition is the pedagogical practice, namely, the ways of how L2 sounds are taught in the process of L2 learning (for instance, how L2 instructors explicitly associate L2 sounds to L1 sounds may facilitate or impede L2 acquisition). For example, while the dentals were found to pose most difficulty in both perceptual tasks, the learners produced frication duration similar to that of the native speakers in the case of aspirated dentals, but not for the other aspirated non-dentals. Therefore, the difficulty in perception did not impede the production of the frication duration in the aspirated dentals of either group of learners, but not along other acoustical dimensions. As mentioned above, the similarity between Mandarin dentals and English stop-fricative sequences and also the methods utilized in Chinese as a foreign language pedagogy may help explain the better production of frication in the aspirated dentals by the learners. In this sense, this finding does not necessarily lend support to the findings reported in Goto (1971) and Sheldon and Strange (1982), namely that L2 production may precede perception.

6. CONCLUSION

In this study, we examined the perception and production of six Mandarin affricates by American L2 learners. It was shown that the learners performed equally well as the native speakers in the discrimination task, but they did not perform as well as the native speakers in the identification tasks, even though the identification difficulties were not consonant-specific. The results of the production task showed that the L2 learners did not produce affricates as well as the native speakers within various acoustical parameters. It was also found that the intermediate learners approximated the native speakers over the beginning learners in the production of some acoustical parameters, thus indicating some learning effects. The fact that the learners produced the aspiration in dentals better than in non-dentals highlights the potential

117

influence of Chinese as a Foreign Language pedagogy on L2 production and perception. Our findings point to the need for future studies to take into account the synergistic interaction between L1, L2, and language teaching pedagogy in instructed second language acquisition.

This study has some limitations. On the one hand, the sample size of each group was small, and the groups were not gender balanced. In addition, the L2 proficiency of the learners was only measured by the duration of language learning. For future studies, the group sample size should be increased, and the genders of the participants should be more balanced. Further, the level of the proficiency of the L2 learners should be measured more objectively.

REFERENCES

Bates, Douglas, Martin Maechler and Ben Bolker. 2011. lme4, R. Package version 0.999375-38.

Bent, Tessa. 2005. *Perception and Production of Non-native Prosodic Categories.* Chicago, IL: Northwestern University dissertation.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, eds. By Winfrend Strange, pp. 171-204. Baltimore: MD: York Press.

Best, Catherine T. and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second Language Speech Learning: In Honor of James Flege*, eds. by Ocke-Schwen Bohn and Murray J. Munro, pp.13-34. Amsterdam: John Benjamins.

Duanmu, San. 2000. *The Phonology of Standard Chinese.* New York: Oxford University Press.

Eckman, Fred. R. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27:315-330.

Eckman, Fred. R. 1991. The structural conformity hypothesis and the acquisition clusters in the interlanguage ESL learners. *Studies in Second Language Acquisition* 13:23-41.

Flege, James E. 1993. Production and perception of a novel, second-language phonetic contrast. *J. Acoust. Soc. Am.* 93:1589-1608.

Flege, James E. 1995. Second-language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience*, ed. by Winfrey Strange, pp.233-277. Timonium, MD: York Press.

Flege, James E., Ocke-Schwen Bohn and Sunyoung Jang. 1997. The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics* 25:437-470.

Flege, James F., Ian R. A. Mackay and Diane Meador. 1999. Native Italian speakers' production and perception of English vowels. *J. Acoust. Soc. Am.*106:2973-2987.

Flege, James E., Carlo Schirru and Ian R. A. MacKay. 2003. Interaction between the native and second language. *Speech Communication* 40:467-491.

Gass, Susan, Jennifer Behney and Luke Plonsky. 2013. *Second Language Acquisition: An Introductory Course*. New York/London: Routledge.

Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia* 9, 3:317-323.

Hao, Yen-Chen. 2012. The Effect of L2 Experience on Second Language Acquisition of Mandarin Consonants, Vowels and Tones. Bloomington, IN: Indiana University at Bloomington dissertation.

Huang, Ting. 2013. *The Acquisition of Mandarin Palatal Consonants: From the Perspective of Spanish-speaking Learners*. Hong Kong, China: The Chinese University of Hong Kong master thesis.

Johnson, Keith. 2002. *Acoustic and Auditory Phonetics*. Oxford: Blackwell.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Lai, Yi-Hsiu. 2009. Asymmetry in Mandarin affricate perception by learners of Mandarin Chinese. *Language and Cognitive Processes* 27, 7/8:1265-1285.

Li, Fangfang, Jan Edwards and Mary E. Beckman. 2007. Spectral measures for sibilant fricatives of English, Japanese and Mandarin Chinese. In *Proceedings of the XVIth International Congress of Phonetic Sciences, 6 -10 August, 2007*, Saarbruecken, Germany.

Li, Chunxiang. 2012. *The Error Analysis of Chinese Affricate and Fricative Acquired as a Second Language by Thai Learners*. Xiamen, China: Xiamen University master thesis.

Lin, Yen-Hwei. 2007. *The Sounds of Chinese*. Cambridge: Cambridge University Press.

Lin, Yigao. 2006. *An Experimental Study on the Indonesian Learners' Acquisition of Stops and Affricates of Standard Chinese*. MA thesis. Guangzhou, China: Jinan University.

Lin, Yigao and Wang Gongping. 2005. An experimental study on the acquisition of stops and affricates in Chinese by Indonesian learners. *Language Teaching and Research* 04:59-65.

Liu, Jiang and Allard Jongman. 2012. American Chinese learners' acquisition of L2 Chinese affricates /ts/ and /tsh/. In *Proceedings of Meetings on Acoustics*, Vol 18, 060005 (2013).

Liu, Xun. 2003. *New Practical Chinese Reader*. Beijing: Beijing Language and Culture Press.

Liu, Yuehua, Tao-Chung Yao, Liangyan Ge, Yen-fen Chen, Nyang-Ping Bi, Xiaojun Wang and Yaohua Shi. 2008. *Integrated Chinese*. Level I Part I. 3rd Edition. Boston: Cheng and Tsui Company, Inc.

Major, Roy C. 2001. *Foreign Accent*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mei, Li. 2011. Perceptual assimilation and differentiation of Mandarin affricates by Thai learners. *World Chinese Language Teaching* 25, 2:276-288.

Mirman, Daniel. 2014. *Growth Curve Analysis and Visualization Using R*. Boca Raton, FL: Chapman and Hall.

Neufeld G. 1988. Phonological asymmetry in second-language learning and performance. *Language Learning* 38, 4:531-560.

Piske, Thorsten, Ian R. A MacKay and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29:191-215.

Sheldon, Amy and Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3:243-261.

Si, Qiuxue. 2011. *The Acquisition of the Fricatives and Affricates in Mandarin by Elementary Level Learners*. Singapore: National University of Singapore master thesis.

van de Weijer, Jeroen and Frans Hinskens. 2004. On segmental complexity: Affricates and patterns of segmental modification in consonant inventories. *Linguistics in the Netherlands* 2004:217-228.

Weinreich, Uriel. 1953. *Language in Contact: Findings and Problems*. The Hague: Mouton.

Wu, Sue-mei, Yueming Yu, Yanhui Zhang and Weizhong Tian. 2010. *Chinese Link*. Pearson. Jo.

Zhang, Hang. 2007. *The Second Language Acquisition of Mandarin Chinese tones by English, Japanese and Korean speakers*. Chapel Hill, NC: The University of North Carolina at Chapel Hill dissertation.

Zhang, Hang. 2013. *The Second Language Acquisition of Mandarin Chinese Tones by English, Japanese and Korean Speakers*. Unpublished Ph. D dissertation. The University of North Carolina at Chapel Hill.

Chunsheng Yang
Department of Literatures, Cultures, and Languages
The University of Connecticut
Oak Hall 207, Fairfield Way 365 U-1057
Storrs, CT 06269
Chunsheng.yang@uconn.edu

Alan C. L. Yu
Department of Linguistics
The University of Chicago
Chicago, IL 60637
aclyu@uchicago.edu

Chunsheng Yang; Alan C. L. Yu

以英文為母語的二語學習者對中文破擦音的習得

楊春生 [1]、余梓麟 [2]
[1] 康涅狄格大學
[2] 芝加哥大學

第二語言中破擦音的習得經常給二語學習者造成困難。本文基於一個發音實驗和兩個感知實驗研究以英文為母語的初、中級學習者對六個中文破擦音的發音與感知。感知實驗顯示二語學習者在區分實驗中與母語持有者無差異，但在識別實驗中二語學習者表現不及母語持有者，各破擦音之間的習得並沒有差異。發生實驗揭示二語學習者的破擦音在多個聲學指標上異於母語者。本文結尾討論了二語發音和感知之間的複雜關係，並探討了母語、二語和二語教學法對課堂二語習得的影響。

關鍵字：破擦音、中文、同化、二語習得、顯著性