SECOND LANGUAGE LEARNERS' ACQUISITION OF MANDARIN FRICATIVE CONTRASTS: AN ULTRASOUND STUDY

Madeleine Oakley

North Carolina State University
mjoakley@ncsu.edu

ABSTRACT

This study examines how second language learners transfer articulatory and acoustic categories to produce non-native fricative contrasts. It is hypothesized that L1 English-L2 Mandarin learners will have difficulty producing the Mandarin /ʂ/-/ɕ/ contrast because both sounds are perceptually assimilated to English /ʃ/ [1]. Results from multi-taper spectra show that learners who produce Mandarin /ʂ/ and /ɕ/ similarly also tend to produce them similarly to English /ʃ/. However, ultrasound tongue imaging results reveal that these same learners do produce a difference in tongue position between Mandarin /ʂ/ and /ɕ/. Furthermore, L1 English-L2 Mandarin learners produce all English fricatives with a higher tongue body than Mandarin fricatives, showing learners are not transferring articulatory gestures to produce these fricatives. The results of this study show that learners are producing the non-native fricative contrast articulatorily, although this finding is obscured when only observing the acoustic data.

Keywords: L2 phonetics, articulation, ultrasound, fricatives, sibilants

1. INTRODUCTION

The primary objective of this study is to examine how learners produce the articulatory contrast between non-native consonant that are predicted to be perceived as phonetically similar to the same L1 consonant category. When acquiring a non-native contrast, it is predicted that learners will assimilate non-native phones to L1 sound categories if the two sounds are perceived as phonetically similar, but will create new categories for L2 sounds if that sound is not perceived as phonetically similar to any L1 sound [2]. If the L2 phone is perceived as similar to an L1 category, learners are expected to produce the L2 phone as if it were an L1 phone [2, 3]. However, learners often use different strategies than native speakers to produce non-native contrasts, which may obscure acquisition of a contrast [4]. In a study examining the perception and production of second language Korean fricatives, learners did not produce a target-like acoustic difference between L2 Korean alveolar fricatives but did perceive the contrast, suggesting the acoustic results obscure information about whether learners have acquired the contrast [5]. Learners may be using different articulatory strategies to produce a contrast than native speakers, which can lead to non-target like, or 'merged', acoustic productions. Thus, there is a possibility that learners are not transferring L1 categories to produce L2 phones, but only using acoustic data and native-like comparisons misinforms theories of category formation.

This study examines both the articulatory and acoustic properties of L1 English-L2 Mandarin learners' fricative productions. Mandarin has a three-way contrast between the fricatives /s/, /ʂ/, and /ɕ/ [6, 7]. It has been shown that L1 English learners of Mandarin have difficulty perceiving the difference between the Mandarin fricatives /ʂ/ and /ɕ/, and it is predicted that this is because they are assimilating the two Mandarin fricatives to their L1 English fricative /ʃ/ [1]. Because learners are predicted to have difficulty producing the difference between a contrast if they are perceived as the same L1 phoneme [2, 3], learners are expected to have difficulty producing the contrast between /ʂ/ and /ɕ/. Mandarin /s/, on the other hand, is predicted to be perceived as similar to English /s/, and thus produced as such. This study uses ultrasound tongue imaging to examine whether learners transfer L1 English acoustic or articulatory categories to produce Mandarin /ʂ/ and /ɕ/, and whether this differs from patterns of transfer for Mandarin /s/.

2. METHODS

2.1. Participants

8 L1 English-L2 Mandarin learners (2 male, 6 female, mean age=22) participated in a production task in English and in Mandarin. All learners were enrolled in intermediate Chinese language courses at a university in the U.S. at the time of

study. 2 L1 Mandarin speakers also completed the Mandarin production task (1 male, 1 female, mean age=25.5). Native speakers reported only speaking the Mandarin dialect of Chinese and speak Chinese on a regular basis with family in China and friends in the U.S.

2.2. Materials

The target Mandarin fricatives /s, ʂ, ɕ/ were presented in the onset of 6 Mandarin words, repeated twice for a total of 12 tokens of each target sound. The words were embedded in the carrier phrase [wə tɕei də __ xən xau], meaning "I think __ (is) very good". Majority of the words were bisyllabic words of the shape CVCV. Following vowel quality and tone were controlled for to the extent possible, while only using lexical items familiar to intermediate learners. All target items were chosen from a vocabulary list given to them in class. The target English fricatives /s, ʃ/ were also presented in the onset position of 6 bisyllabic English words, repeated twice for a total of 12 tokens of each target sound. The words were embedded in the carrier phrase "give me a __ again". Following vowel context was controlled for across the English fricatives, and were similar to the environment of the Mandarin following vowel contexts.

2.3. Procedure

Participants wore a head stabilization headset manufactured by Articulate Assistant Advanced (AAA) [8], attached to an ultrasound probe and a video camera. Participants also wore a head-mounted microphone. All acoustic and articulatory data were recorded using AAA. Learners completed the Mandarin task first, followed by the English task. The Mandarin words were written in Chinese characters.

2.4. Analysis

First, each speakers' audio prompts were aligned to transcriptions using MFA [9]. The aligned textgrids were hand-corrected in Praat and re-uploaded to AAA for the articulatory analysis. For the acoustic analysis, a Praat script was used to extract all the fricatives from the target phrases. Spectra were measured using multi-taper spectral analysis in the middle 20 ms portion of each fricative with the spectRum package for R [10, 11]. Spectral center of gravity and spectral speak were measured for each fricative, following from work that argues these acoustic measures correlate to the Mandarin three-way fricative contrast [12, 7].

For the articulatory analysis, splines were semi-automatically fitted to the ultrasound images for each speaker. Splines were extracted out of AAA at the acoustic midpoint using polar coordinates. The midpoint is chosen for analysis to avoid the coarticulatory effects of preceding or following vowel context as much as possible. For the L1 Mandarin speakers, the tongue positions of Mandarin fricatives /s, ʂ, ɕ/ were compared using SSANOVAs in order to confirm how speakers produce the articulatory contrast between the target segments. Next, the learner productions of Mandarin /s, ʂ, ɕ/ were compared within each speaker to see whether learners were using the same articulatory strategies as native speakers to produce this contrast. Finally, each learners' productions of Mandarin /s, ʂ, ɕ/ were compared to their English productions of /s, ʃ/ using SSANOVAs to determine whether learners were transferring L1 articulatory gestures to produce the L2 fricatives. All analyses were conducted within participants (rather than across participants) in order to show how each speaker is producing the contrast between Mandarin fricatives, and to examine whether learners were transferring their L1 categories to produce Mandarin fricatives.

3. RESULTS

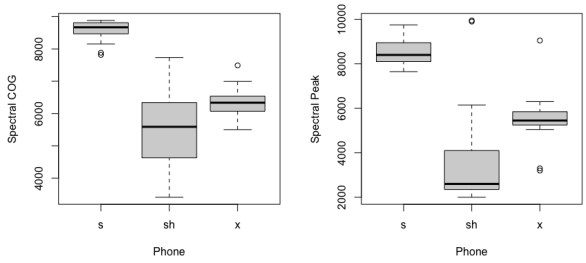
3.1. Acoustic Results

First, separate one-way ANOVAs for each native speaker with spectral COG as the dependent variable confirms that there is a significant difference in COG between the target fricatives (results for NM1 are shown in Figure 1a; F=29.38, p=3.77e-10***).

Similarly for spectral peak, separate ANOVAs confirm that there is a significant difference in spectral peak for native speaker productions' of Mandarin fricatives (F=51.08; p=1e-14***, shown in Figure 1b). Post-hoc Tukey HSD tests confirm that COG and spectral peak are different for all three fricatives.

Next, learner acoustic productions of the Mandarin fricative contrast were compared using one-way ANOVAs. Again, results are shown within-participant to confirm how learners are producing the non-native contrast. One learner's results are shown here as representative of group results, and other learner results are summarized.

Results show that all learners produce a difference in COG between the Mandarin fricatives. However, Tukey HSD tests confirm that while all learners produce a difference in COG between Mandarin

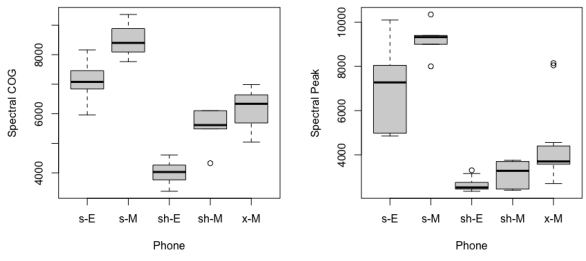


(a) Center of Gravity (Hz) **(b)** Spectral Peak (Hz)

Figure 1: Native speaker NM1's spectral measurements of Mandarin fricatives; Mandarin fricatives /ʂ/ and /ɕ/ are represented by their Pinyin orthography 'sh' and 'x'

/s/ and /ʂ/ and Mandarin /s/ and /ɕ/, only half of the learners produce a difference in COG between Mandarin /ʂ/ and /ɕ/. Results for learner LM4, who does not produce a difference between Mandarin /ʂ/ and /ɕ/, are presented in Figure 2a (F=92.08, p<2e-16***).

For spectral peak, separate one-way ANOVAs for each learner show that learners do produce a difference in spectral peak for the Mandarin fricatives. Post-hoc Tukey HSD tests show that all learners produce a difference between /s/ and /ɕ/, and again, half of the learners produce a difference between /ʂ/ and /ɕ/. Results for LM4, who does not produce a difference in spectral peak between Mandarin /ʂ/ and /ɕ/, are presented in Figure 2b; F=38.77, p=1.78e-14***). All learners except one (LM1) produce a difference between Mandarin /s/ and /ʂ/.



(a) Center of Gravity (Hz) **(b)** Spectral Peak (Hz)

Figure 2: LM4's spectral measurements of English and Mandarin fricatives

Finally, each learners' productions of their Mandarin fricatives were compared to their English fricatives using separate one-way ANVOAs for each acoustic measurement with fricative category as the dependent variable. There is a significant difference in COG between fricatives for all learners. Tukey HSD tests show that most learners do not produce a difference in COG between English /s/ and Mandarin /s/, or between English /ʃ/ and Mandarin /ʂ/. LM4 (Figure 2a) is the only speaker who produces a difference in COG between these fricatives. All learners produce a difference in COG between English /s/ and Mandarin /ʂ/, and English /s/ and Mandarin /ɕ/. Half of the learners produce a difference in COG between English /ʃ/ and Mandarin /ɕ/.

All learners produce a difference in spectral peak between fricatives as well. Post-hoc Tukey HSD tests show that no learners (except LM4, Figure 2b) produce a difference in spectral peak between English /s/ and Mandarin /s/. No learners produce a difference between English /ʃ/ and Mandarin /ʂ/. Half of the learners produce a difference English /ʃ/ and Mandarin /ɕ/. All learners do produce a difference between English /s/ and Mandarin /ʂ/, and English /s/ and Mandarin /ɕ/.

To summarize the acoustic results: native speakers produce a difference in COG and spectral peak between the three Mandarin fricatives /s/, /ʂ/, and /ɕ/. L1 English-L2 Mandarin learners do produce a difference in COG and spectral peak between the Mandarin fricatives /s/ and /ʂ/, and /s/ and /ɕ/, but only half produce a difference between Mandarin /ʂ/ and /ɕ/. Turning to the cross-language results, most learners produce Mandarin /s/ with the same COG and spectral peak as English /s/. Most learners also produce Mandarin /ʂ/ with the same COG and spectral peak as English /ʃ/. About half of the learners produce a difference in COG and spectral peak between English /ʃ/ and Mandarin /ɕ/. Many of these learners produce the Mandarin fricatives /ʂ/ and /ɕ/ with the same COG and spectral peak.

3.2. Ultrasound results

Native speaker NM1's SSANOVA results for the production of Mandarin fricatives are shown in Figure 3. There is a significant difference in native speakers' productions of the Mandarin fricatives at the constriction location.

Next, learner productions of the Mandarin fricatives were compared (again, within speaker) using SSANOVAs. Results for LM5's productions of Mandarin and English fricatives are shown in Figure 4a. As can be seen, this learner does not produce a difference in tongue position for Mandarin /s/ and /ʂ/. Half of the learners in this study do not produce Mandarin /s/ and /ʂ/ with significantly different tongue positions (similar to LM5), while half do produce a difference in tongue position (see Figure 4b for learner LM4's

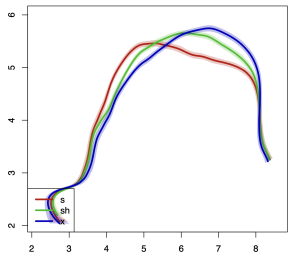
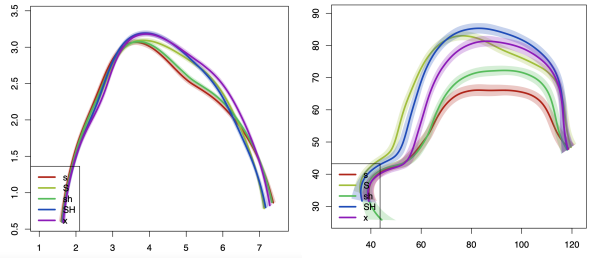


Figure 3: NM1's SSANOVA results (tongue tip to the right) for the production of Mandarin fricatives

SSANOVA results). However, Mandarin /ɕ/ is produced with a different tongue position than either of the other two Mandarin fricatives for learners. All learners in this study do produce a significant difference between Mandarin /s/ and /ɕ/ and Mandarin /ʂ/ and /ɕ/ at least along the tongue front.



(a) LM5's SSANOVA results **(b)** LM4's SSANOVA results

Figure 4: Learner SSANOVA results (tongue tip to the right) for the production of Mandarin and English fricatives (note that English fricatives /s/ and /ʃ/ are indicated as capital letters S and SH while Mandarin fricatives /s/, /ʂ/ and /ɕ/ are indicated by their Pinyin orthography s, sh, and x)

Turning to the cross-language results, all learners tend to produce a significant difference in tongue position between Mandarin /s/ and English /s/, and Mandarin /ʂ/ and English /ʃ/. Again, this is shown in Figures 4a and 4b. Interestingly, most learners produce their English fricatives with a higher position than their Mandarin fricatives. Finally, all learners do produce a difference between Mandarin /ɕ/ and English /ʃ/. This is surprising given the acoustic findings that half of the learners produce Mandarin /ɕ/ with similar spectral measures to English /ʃ/.

4. DISCUSSION

This study examines how L1 English-L2 Mandarin learners acquire the non-native contrast between the Mandarin fricatives /s/-/ʂ/-/ɕ/, and whether learners transfer acoustic categories from L1 fricatives to produce the non-native fricatives, or articulatory gestures. The results show individual differences in patterns of transfer. Importantly, learners who do not produce an acoustic difference between the Mandarin fricatives /ʂ/-/ɕ/ do produce an articulatory difference between these sounds, and do not transfer L1 articulatory gestures to produce these fricatives.

Half of the learners in this study do not produce a difference in spectral measures between Mandarin /ʂ/ and /ɕ/. These sounds are both acoustically similar to English /ʃ/ (although this varies across learners). This result may lead to the conclusion that the learners who do not produce the difference between these sounds have not acquired the contrast (or assimilated them to the same L1 category). However, articulatory results show that /ʂ/ and /ɕ/ are produced with different tongue positions by all speakers, including those who do not produce these fricatives with different spectral measures.

The lack of spectral difference between L2 Mandarin /ʂ/-/ɕ/ for many learners may be caused by learners producing this contrast as a difference in tongue height rather than constriction location. The spectral measures included in this study are correlated to the size of the front cavity between the lingual constriction and the lip opening, which does not reflect how these learners are producing the contrast. Importantly, native Mandarin speakers do produce a difference in COG and spectral peak between /s/ /ʂ/ and /ɕ/. The results presented here show that there are a variety of individual strategies learners use to produce the contrast, but these are often different articulatory strategies than native speakers. Only observing the acoustic results obscures the fact that learners do appear to have the contrast.

Finally, it is not the case that learners transfer L1 articulatory gestures to produce L2 Mandarin fricatives. There is no evidence of articulatory transfer, even for similar phones /s/ and /s/. Again, this was not seen in the acoustic results.

The results of this study highlight the fact learners often use different strategies to produce L2 contrasts from native speakers, and including both acoustic and articulatory data in the study of category formation can reveal covert contrasts. This study also shows that learners do not transfer L1 articulatory gestures to produce L2 fricatives, which suggests L1 transfer may not be as straightforward as assumed from models based solely on acoustic data.

5. REFERENCES

[1] X. Wang and J. Chen, "English speakers' perception of mandarin consonants: The effect of phonetic distances and l2 experience," *ICPhS 2019 Proceedings*, pp. 1–5, 2019.

[2] J. E. Flege, "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification," *Journal of phonetics*, vol. 15, no. 1, pp. 47–65, 1987.

[3] ——, "Origins and development of the speech learning model," *Retrieved December*, vol. 13, p. 2005, 2005.

[4] R. McAllister, "Strategies for realization of l2-categories," *Language experience in second language speech learning: In honor of James Emil Flege. John Benjamins Publishing*, vol. 17, pp. 153–166, 2007.

[5] S. Y. Cheon, "Production of korean fricatives in second language acquisition: Acoustic characteristics," *Korean Linguistics*, vol. 13, no. 1, pp. 17–48, 2006.

[6] S. Duanmu, *The phonology of standard Chinese*. OUP Oxford, 2007.

[7] F. Li, J. Edwards, and M. Beckman, "Spectral measures for sibilant fricatives of english, japanese, and mandarin chinese," in *Proceedings of the 16th International Congress of Phonetic Sciences*. Pirrot, Dudweiler Germany, 2007, pp. 917–920.

[8] AAA, "Articulate instruments ltd.: Articulate assistant advanced user guide: Version 218.04." 2019.

[9] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, "Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi," in *Proc. Interspeech 2017*, 2017, pp. 498–502.

[10] P. F. Reidy, "spectRum," 2013, r package. [Online]. Available: https://github.com/patrickreidy/spectRum

[11] R Core Team, "R language definition," 2015.

[12] S.-I. Lee, "Spectral analysis of mandarin chinese sibilant fricatives." in *ICPhS*, 2011, pp. 1178–1181.