A Tale of Five Fricatives: Consonantal Contrast in Heritage Speakers of Mandarin

Charles B. Chang, Erin F. Haynes,
Yao Yao, and Russell Rhodes*

1 Introduction

Heritage speakers – that is, speakers who have had exposure to a particular language as a child, but who have shifted to another language for the majority of their communication needs – have begun to draw attention in the field of phonological learning. Au, Jun, Knightly, and Oh have jointly explored the phonological competence of heritage speakers in both their subjects' heritage language and main language. They find that heritage speakers of Spanish and Korean tend to have a phonological advantage over late learners in production of the heritage language, as indicated by acoustic measures such as voice onset time and by holistic perceptual measures such as accent ratings by native speakers (cf. Au *et al*. 2002; Knightly *et al*. 2003; Oh *et al*. 2002, 2003). Of the few studies on heritage language phonology, however, only Godson (2003) explores the neutralization of phonological categories in the heritage language, and only with respect to vowels. Her findings suggest that the Armenian vowels of heritage speakers of Armenian are influenced by their dominant language, English, but that this influence is limited to those Armenian vowels that are close to English vowels and does not necessarily result in the neutralization of contrast.

The present study extends this line of inquiry to consonants by comparing fricative production in heritage speakers of Mandarin to that of native Mandarin speakers and native English speakers who have learned Mandarin as a foreign language. We focus on place contrasts among five voiceless fricatives, in particular one between two post-alveolar Mandarin fricatives, retroflex /ʂ/ and alveolo-palatal /ɕ/. Figure 1 shows that in comparison to /ʂ/, the area of contact for /ɕ/ is slightly more forward, going right up to the incisors, as well as significantly wider, extending sideways onto the molars and

*For helpful comments and discussion at multiple stages of this research, we would like to thank Sharon Inkelas, Keith Johnson, participants in a fall 2007 UC Berkeley seminar on phonological learning, and audience members at PLC 32. We are also grateful to all the Mandarin speakers and learners who took part in this study. This work was supported in part by a National Science Foundation Graduate Research Fellowship to the first author. Any errors are our own.

much farther inwards onto the hard palate. This pattern of contact results in a smaller front cavity and narrower channel area for /ɕ/ in comparison to /ʂ/ – both properties which affect the quality of the noise in /ɕ/ vs. /ʂ/.

Figure 1: Palatograms of Speaker 1's fricatives in /ʂa^{51}/ 'suddenly' (left) and /ɕa^{51}/ 'below' (right).

In addition to examining the realization of this Mandarin contrast, this study investigates whether heritage speakers distinguish between these post-alveolar fricatives and the post-alveolar fricative of their dominant language (namely, the English palato-alveolar fricative /ʃ/), as well as whether they distinguish between the Mandarin alveolar fricative /s/ and the English alveolar fricative /s/. These are all pairs of consonants that, due to their high degree of phonetic similarity, stand to undergo "equivalence classification" (Flege 1987) and thereby become indistinguishable from each other. Using acoustic measures of place of articulation, we thus examine both the question of whether heritage speakers maintain consonantal contrasts in Mandarin, as well as the question of whether they maintain contrasts between Mandarin consonants and similar English consonants.

2 Methods

2.1 Participants

Eighteen Mandarin speakers and learners participated in this study. Five were native Mandarin speakers who were born and educated in a Mandarin-speaking country; eight were heritage speakers of Mandarin who either were born in or came to the U.S. before the age of 10; and five were native English speakers who were born and educated in the U.S. and had learned Mandarin as a foreign language in high school or college. Speakers were as-signed to these groups, as well as rank-ordered within them, based on a detailed questionnaire about their language background, current language use, and comprehension of Mandarin in formal and informal situations. Participants ranged in age from 18 to 40 years old, and none reported any history of speech or hearing impairments.

2.2 Stimuli

Participants were presented with 62 Mandarin words and phrases and 35 English words in random order via individual index cards. English words were written in English orthography, and Mandarin words were written in Mandarin orthography (traditional or simplified characters) and romanization (pinyin and/or BoPoMoFo). Critical stimuli contained one of the two Mandarin post-alveolar fricatives, /ʂ̻/ and /ɕ/, or one of three other fricatives – English /ʃ/, Mandarin /s/, and English /s/. These fricatives appeared prevocalically in ten monosyllabic Mandarin words and five monosyllabic English words (see the appendix for a full list).

2.3 Recording

Recording was done in a sound-proof booth at 48 kHz and 16 bps. The equipment used was either a Marantz PMD660 solid-state recorder with an AKG C420 head-mounted condenser microphone, or an M-AUDIO Mobile-Pre USB preamp audio interface with an AKG C520 head-mounted condenser microphone. Stimuli were recorded in eight blocks (four blocks of Mandarin and four blocks of English), resulting in a total of four tokens of each item. Blocks were grouped by language such that participants completed all blocks in one language before moving on to blocks in the other language, with the order of the languages (Mandarin-English or English-Mandarin) balanced across participants. Data for English /s/ were unable to be obtained for two speakers (Speakers 11 and 18), while data for Mandarin /s/ were unable to be obtained for one speaker (Speaker 11).

2.4 Acoustic Analysis

All measurements were taken by hand in Praat (Boersma and Weenink 2008). Peak amplitude frequency (PAF) and centroid frequency (Ladefoged 2005) were measured over a spectrum of the middle 100 ms of the fricative. The transitional first (F1), second (F2), and third (F3) formants from the fricative to the following vowel were also measured over the first 20 ms of the vowel.

To ensure that the measurements taken were reliable, 25% of each of the

PAF, centroid, F1, F2, and F3 measurements were double-checked by a second researcher. Any discrepancy between the two researchers' measurements in excess of 100 Hz was checked again by a third researcher, resulting in 19% of the total number of measurement checks being triple-checked. Final calculations of the differences between researchers' measurements revealed an average difference of 11 Hz in PAF measurements (88% less than 25 Hz apart), 39 Hz in centroid measurements (39% less than 25 Hz apart), 18 Hz in F1 measurements (77% less than 25 Hz apart), 16 Hz in F2 measurements (81% less than 25 Hz apart), and 19 Hz in F3 measurements (75% less than 25 Hz apart). If after a third measurement there still remained a discrepancy between different researchers' measurements of greater than 100 Hz, all of these measurements were discarded; however, this resulted in the disposal of less than 1% of the total number of measurements.

3 Results

Apart from showing /ɕ/ to be the most "palatalized", formant transitions do not differentiate the fricatives very clearly, so we concentrate here on data from PAF and centroid frequency. The differences between Mandarin and English fricatives in terms of centroid frequency are summarized in Table 1 (Mandarin figures averaged from Svantesson 1986, English figures from Jongman *et al*. 2000).[1] As seen here, the average centroid for English /s/ is slightly higher than that of Mandarin /s/. As for the post-alveolar fricatives, Mandarin /ɕ/ has the highest centroid, followed by English /ʃ/ and Mandarin /ʂ/. Thus, no two of these fricatives are the same with respect to centroid.

Mandarin category		Centroid
alveolar	/s/	6006
retroflex	/ʂ/	3585
alveolo-palatal	/ɕ/	5381

English category		Centroid
alveolar	/s, z/	6133
palato-alveolar	/ʃ, ʒ/	4229

Table 1: Native centroid targets (in Hz) for Mandarin and English fricatives.

Graphs of mean PAF and centroid for all fricatives and speakers are given in Figures 1-4 below, separated by gender (Speakers 1-5 are native Mandarin speakers; 6-13, heritage speakers; and 14-18, late learners).

[1]Note that the average centroids for /s/ and /ʃ/ are likely to be slightly higher than the figures given in Table 1, since these are averages that include the corresponding voiced fricatives (whose centroids will be drawn down by the lower frequencies of f_0).

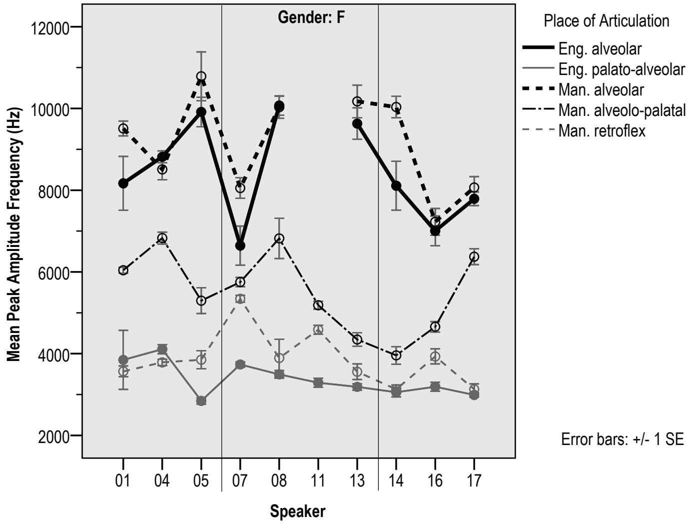

Figure 1: Mean peak amplitude frequency by fricative (female speakers).

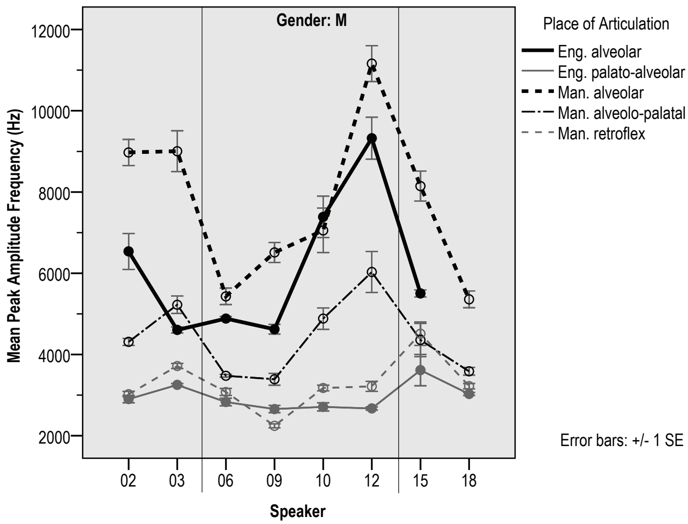

Figure 2: Mean peak amplitude frequency by fricative (male speakers).

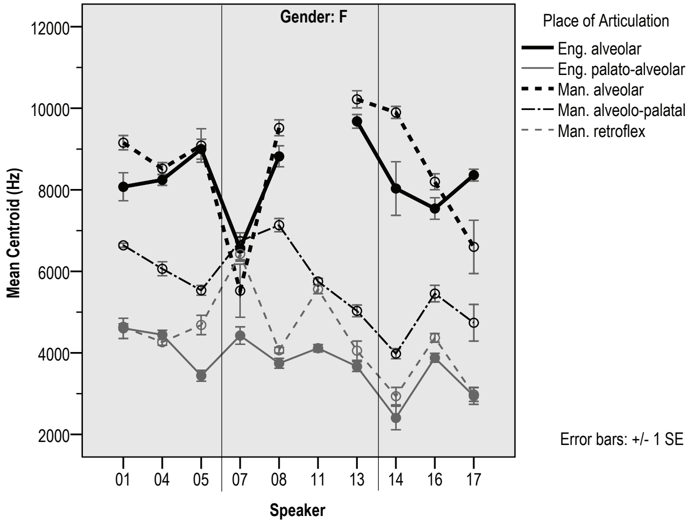

Figure 3: Mean centroid frequency by fricative (female speakers).

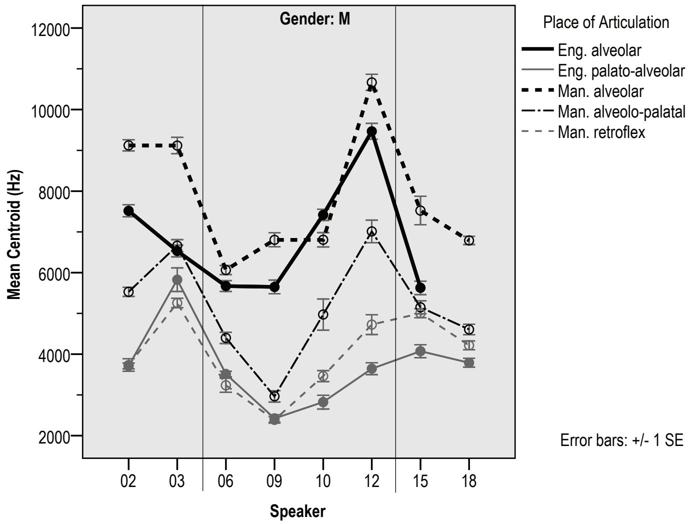

Figure 4: Mean centroid frequency by fricative (male speakers).

Planned comparisons using the Wilcoxon matched pairs signed-rank test reveal four main patterns. First, the Mandarin post-alveolar fricatives /ʂ/ and /ɕ/ are distinguished by nearly everyone (cf. Table 2 below). For 17 out of 18 speakers, the difference between /ʂ/ and /ɕ/ is statistically significant at $p < .05$ with respect to PAF or centroid and, in the majority of cases, highly significant on both measures. The one speaker who fails to distinguish these fricatives on either (Speaker 15) is a late learner, as one might expect.[2]

Speaker	/ʂ/ vs. /ɕ/	/ɕ/ vs. /ʃ/	/ʂ/ vs. /ʃ/	Man. /s/ vs. Eng. /s/
1	**	(*)	n.s.	n.s.
2	**	*	n.s.	**
3	**	(*)	/	**
4	**	*	n.s.	n.s.
5	*	*	*	n.s.
6	**	*	n.s.	(*)
7	/	*	*	/
8	**	*	n.s.	n.s.
9	**	*	/	**
10	*	*	/	/
11	/	*	*	--
12	**	*	(*)	/
13	*	*	n.s.	n.s.
14	*	*	n.s.	*
15	n.s.	/	/	**
16	(*)	*	*	n.s.
17	**	*	n.s.	/
18	*	*	n.s.	--

Table 2: Distinctions made between fricatives, by speaker and contrast.

Speakers 1-5: native Mandarin speakers; 6-13: heritage speakers; 14-18: late learners. *: $p < .05$, **: $p < .01$ on both PAF and centroid; (*): $p < .05$ on PAF or centroid and approaching significance on the other; /: $p < .05$ on PAF or centroid but not significant on the other; n.s.: *p* significant on neither PAF nor centroid; --: data unavailable.

Second, all speakers distinguish Mandarin /ɕ/ and English /ʃ/. For 15

[2]Incidentally, no speakers show dialectal neutralization of Mandarin post-alveolar /ʂ/ with alveolar /s/; rather, they distinguish both /ʂ/ and /ɕ/ from /s/ in PAF or centroid (in fact, 16 out of the 17 speakers for whom data on Mandarin /s/ is available distinguish each pair of fricatives along both of these dimensions).

speakers, including all of the heritage speakers, the difference between /ɕ/ and /ʃ/ is significant with respect to both PAF and centroid; for the other three, the difference is significant with respect to only one of these measures (though approaching significance on the other measure in two cases).

On the other hand, only half the speaker pool distinguishes Mandarin /ʂ̻/ and English /ʃ/. While nine speakers show a significant difference between the two fricatives on PAF and/or centroid, the other nine do not. However, the nine that do distinguish them are not evenly distributed across the three speaker groups; instead, the majority of these "distinguishers" are clustered in the heritage speaker group, with the result that the majority of both native Mandarin speakers and late Mandarin learners are not found to distinguish /ʂ̻/ vs. /ʃ/, whereas the majority of heritage speakers are.

Finally, results for the alveolar fricatives are similar: 10 out of the 16 speakers for whom there is data on both Mandarin /s/ and English /s/ distinguish the two. Again these "distinguishers" are not evenly distributed across speaker groups, but are clustered in the heritage speaker group as well as the late learner group, such that the majority of native Mandarin speakers are not found to distinguish Mandarin /s/ vs. English /s/, while the majority of heritage speakers and late learners are. Note that the group that fails to distinguish the two fricatives includes only female speakers; all male speakers distinguish the two on one or both acoustic dimensions, the vast majority (six out of seven) producing Mandarin /s/ with a higher PAF and centroid frequency than English /s/. This result is in contrast to both the predictions of Table 1 and the results of Li *et al*. (2007), who found instead that English /s/ was produced by (presumably monolingual) English speakers with higher centroid values than those of Mandarin /s/ produced by (presumably monolingual) Mandarin speakers.

4 Discussion and Conclusion

To summarize, we collected productions of Mandarin and English by native speakers, heritage speakers, and late learners of Mandarin and found that all or almost all distinguish Mandarin /ʂ̻/ vs. /ɕ/, as well as Mandarin /ɕ/ vs. English /ʃ/. However, only about half distinguish Mandarin /ʂ̻/ vs. English /ʃ/ or Mandarin /s/ vs. English /s/, with the majority of heritage speakers falling into this group of "distinguishers" in both cases. These results indicate, first, that heritage speakers, in addition to most late learners, do not have much trouble with the Mandarin post-alveolar contrast. Second, they suggest that while native speakers and late learners of Mandarin tend to merge similar Mandarin and English sounds, heritage speakers tend to keep them apart.

There are two possible (though not mutually exclusive) explanations for why heritage speakers seem to do better at maintaining contrast between similar sounds in two languages. First, early exposure to both languages might simply make heritage speakers better able to hit close, but not identical targets accurately. Alternatively, it may be that when similar categories are acquired early, they interact with each other in a shared phonological system and are dissimilated or “polarized” (cf. Laeufer 1997). Our current data cannot conclusively distinguish between these two hypotheses, but the fact that the size of PAF and centroid differences between categories (e.g. PAF of /ʂ/ – PAF of /ʃ/) is not correlated with speaker rank or group, even when speakers are separated by gender, suggests that the former hypothesis is probably closer to the truth. It does not appear to be the case that the phonetic distance between categories increases for heritage speakers in particular.

Finally, we are careful to note that the lack of PAF or centroid differences only suggests that speakers are merging the articulations of different categories. One would need detailed articulatory data (e.g. from ultrasound) to be able to conclude definitively that the articulations have in fact become identical for these speakers. Furthermore, it is not clear what category the “merger” speakers merge towards, although it would stand to reason that they would merge in the direction of their native language.

In short, our findings reveal that not only do heritage speakers achieve better accents in the heritage language as found by Au and colleagues, they also appear better able to maintain phonological contrast, both between individual categories of the heritage language and between categories of the heritage language and similar categories in the dominant language.

Appendix

MANDARIN		ENGLISH
retroflex /ʂ/	alveolo-palatal /ɕ/	palato-alveolar /ʃ/
沙 /ʂa^{55}/ ‘sand’	瞎 /ɕa^{55}/ ‘shrimp’	*shop* /ʃɑp/
啥 /ʂa^{35}/ ‘what’	轄 /ɕa^{35}/ ‘govern’	*shot* /ʃɑt/
傻 /ʂa^{214}/ ‘stupid’	下 /ɕa^{51}/ ‘below’	
煞 /ʂa^{51}/ ‘suddenly’		
alveolar /s/		alveolar /s/
撒 /sa^{55}/ ‘to tell (a lie)’		*sob* /sɑb/
撒 /sa^{214}/ ‘to spread (seeds)’		*sod* /sɑd/
飒 /sa^{51}/ ‘sound of wind’		*sock* /sɑk/

Table 3: Critical stimuli in the production experiment.

References

Au, Terry K., Leah M. Knightly, Sun-Ah Jun, and Janet S. Oh. 2002. Overhearing a language during childhood. *Psychological Science* 13(3):238–243.

Boersma, Paul, and David Weenink. 2008. Praat: doing phonetics by computer. http://www.praat.org.

Flege, James Emil. 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15(1):47–65.

Godson, Linda. 2003. Phonetics of Language Attrition: Vowel Production and Articulatory Setting in the Speech of Western Armenian Heritage Speakers. Doctoral Dissertation, University of California, San Diego.

Jongman, Allard, Ratree Wayland, and Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108(3):1252–1263.

Knightly, Leah M., Sun-Ah Jun, Janet S. Oh, and Terry K. Au. 2003. Production benefits of childhood overhearing. *Journal of the Acoustical Society of America* 114(1):465–474.

Ladefoged, Peter. 2005. *Vowels and Consonants*, 2nd edition. Malden, MA: Blackwell Publishing.

Laeufer, Christiane. 1997. Towards a typology of bilingual phonological systems. In *Second-Language Speech: Structure and Process*, ed. A. R. James and J. Leather, 325–342. Berlin: Mouton de Gruyter.

Li, Fangfang, Jan Edwards, and Mary Beckman. 2007. Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese. In *Proceedings of the 16th International Congress of Phonetic Sciences*, ed. J. Trouvain and W. J. Barry, 917–920.

Oh, Janet S., Terry K. Au, and Sun-Ah Jun. 2002. Benefits of childhood language experience for adult L2 learners' phonology. In *Proceedings of the 26th Annual Boston University Conference on Language Development*, Vol. 2, ed. B. Skarabela *et al*., 464–472. Somerville, MA: Cascadilla Press.

Oh, Janet, Sun-Ah Jun, Leah Knightly, and Terry Au. 2003. Holding on to childhood language memory. *Cognition* 86(3):B53–B64.

Svantesson, Jan-Olof. 1986. Acoustic analysis of Chinese fricatives and affricates. *Journal of Chinese Linguistics* 14(1):53–70.

Department of Linguistics
University of California, Berkeley
1203 Dwinelle Hall #2650
Berkeley, CA 94720-2650
cbchang@berkeley.edu
hayneser@berkeley.edu
yaoyao@berkeley.edu
russell_rhodes@berkeley.edu