The Three Sibilants in Standard Chinese

Fang Hu

Phonetics Lab, Institute of Linguistics, Chinese Academy of Social Sciences
E-mail: fong.hwu@gmail.com

Abstract

This paper examines the place of articulation and articulatory variability for the production of the three sibilants in Standard Chinese. EPG data from 4 speakers and EMA data from 2 speakers show that [s] is anterior to [ʂ] whereas it's meaningless to characterize a distinct place feature of [ɕ]. And the articulatory variability is mainly attributable to the inter-speaker sources.

1 Introduction

Standard Chinese (SC hereafter) is well-known for having a tripartite distinction of place of articulation for its sibilant fricatives and affricates, namely alveolar/dental, postalveolar, and alveolo-palatal. The postalveolars are transcribed and described as retroflex sounds in traditional literature [1], whereas empirical studies have shown that they are actually apical postalveolars, i.e., without a curling-back gesture of tongue tip as in canonic retroflexes in Indian languages [2] [3]. Alveolo-palatals are widely found in Chinese dialects and languages as well as in some Slavic and Finno-Ugric languages. Although the alveolo-palatal is identified as a place of articulation in the IPA, suggesting a somewhat posterior place of articulation to the postalveolar, yet it is not uncontroversial regarding its distinct place of articulation. For instance, it is argued in [2] and [3] that the alveolo-palatals in SC are distinguished from their alveolar/dental and postalveolar cognates in terms of lingual gesture and the overall linguopalatal contact pattern, rather than place of articulation.

Although coronals are most common consonants in the World's languages, a tripartite place distinction for sibilants and affricates is a highly marked case [4]. Given a certain articulatory zone, more place distinctions normally require more preciseness in articulation. It is thus of interest to examine the articulatory variability of these sounds. In an MRI-based cross-linguistic study on the sibilants, [5] indicates that the sibilants [s ʃ] in English and French (two-sibilant languages) show more articulatory variability than the sibilants [s ʂ ʃ] in Chinese and Swedish (three-sibilant languages). [6] reports further that in addition to a strategy of tongue position adjustment, the 7 French speakers also use a strategy of tongue shape adjustment, a [ɕ]-like articulation, in the production of French [ʃ]. That is, the French [ʃ] is quite flexible in lingual articulation. And this also suggests that the place feature itself is not sufficient for capturing fine phonetic details even in a two-sibilant language. Regarding the source of variability, [7] suggests that it is more speaker-dependent than language-dependent. [8] explains further that articulatory variability not only depends on linguistic aspects but also on speakers' morphological and motor constraints, specifically a dome-shaped palate vs. a flat palate.

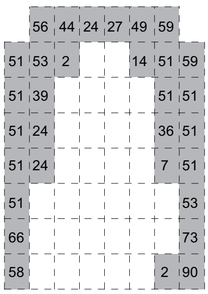
However, pervious studies on SC are mainly based on somehow less natural data such as traditional palatogram/linguogram, X-ray photograph, and MRI and thus have a limited power in explaining the variability. By using the articulatory and acoustic data acquired from more natural experimental setups, this paper examines the issues of place of articulation and articulatory variability for the three sibilants [s ʂ ɕ] in SC.

2 Methodology

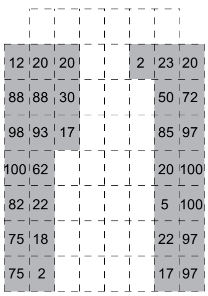
Four speakers, two male and two female, were recorded by using Electronic Palatograph (EPG, the WinEPG system), and two of them, one male and one female, were also recorded by using Electromagnetic Articulograph (EMA, the Carstens AG500 system). The audio signals synchronized with the EMA recording were analyzed acoustically. In addition, palatograms and linguograms were made for one male speaker, as his [s] was purely dental and thus no alveolar contact was observed in the EPG data.

Test material consists of meaningful high level tone monosyllabic words, the target sibilants [s ʂ ɕ] followed by the vowel [i] ([ɿ ʅ] for [s ʂ] respectively), [u] or [a] ([iu ia] for [ɕ]). The test word was placed in a carrier frame composed of four short clauses, with the target occurring in citation, clause-mid, clause-initial and clause-final positions: "X. tʂɤ kə X tsɿ this is X. X tsɿ tɕiɛn dan X is easy. uo lai ʂuə X I am reading X". 5 repetitions were recorded for the EPG study and 8 to 11 repetitions for the EMA study.

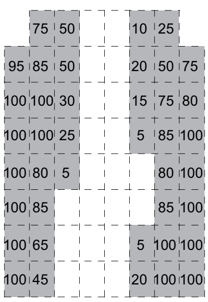
3 EPG Results



[s]



[ʂ]



[ɕ]

Figure 1: *Linguopalatal contact patterns for the three sibilants in SC with the digits indicating the percentage of the contact pooled from all repetitions and all speakers. Data correspond to the midpoint of the frication phase in the citation position.*

Figure 1 shows an overall picture of the linguopalatal contact configurations for the three sibilants in the citation position. The most linguopalatal contact is located in the first (i.e., anterior most) row for [s], suggesting an alveolar place of articulation in general, and in the second or third row for [ʂ], suggesting a postalveolar place of articulation. However, it is difficult to characterize [ɕ] in terms of a distinct place of articulation. Rather, the intensive linguopalatal contact along the whole alveolo-palatal region suggests that the entire anterior part of the tongue is involved in the articulation. And as indexed by the percentage of contact in the figure, [s] is most variable and [ɕ] is least variable, with [ʂ] in-between. It should be noted that the high variability for [s] is partially attributable to MS1, as his [s] is apical (inter-)dental, i.e. without linguopalatal contact along the anterior palatal zone at all, as evidenced by the linguogram and palatogram in Figure 2.

Figure 2: *Palatogram (left) and linguogram (right) of [sa] for MS1.*

The EPG data were quantified by using the alveolar contact anteriority index (CAa), which is a weighted indicator of place of articulation calculated for the first five anterior rows [9], as the articulation of the three sibilants mainly involves the alveolar and palatal zones. In addition, an arithmetical index of the total contact percentage (APT) was calculated for the first six anterior rows, as it is argued above that [ɕ] cannot be characterized by place of articulation. Each index constitutes a dataset of 3 (consonant) × 4 (speaker) × 3 (vowel context) × 4 (clause position).

One-way ANOVAs run on the CAa dataset yielded no significant effects of 'vowel context' ($F = 0.22$, $p = 0.806$) and 'clause position' ($F = 0.28$, $p = 0.84$), respectively. With these two within-speaker factors excluded, a two-way ANOVA run on the dataset of CAa yielded a significant effect of 'consonant' ($F = 297.38$, $p < 0.000$), 'speaker' ($F = 331.74$, $p < 0.000$), and 'consonant' × 'speaker' interaction ($F = 156.25$, $p < 0.000$). The main effects are associated with variations in contact anteriority for [ɕ] > [s] > [ʂ] and for FS1 > FS2 > MS2 > MS1. However, there is no meaningful interpretation for the fact that [ɕ] is the anterior most among the three sibilants. Rather, it obviously contradicts with what is conceptualized in IPA. In short, the results suggest that [s ʂ] are distinguishable from each other in terms of place of articulation, whereas [ɕ] cannot be explained by the place difference at all. Alternatively, APT seems to be a better index in characterizing the three sibilants in SC. A one-way ANOVA yielded a significant effect of 'consonant' ($F = 38.81$, $p < 0.000$). And the results capture the characteristics of linguopalatal contact for all the three sibilants, given that [s] is associated with the smallest percentage of total alveolar and palatal contact, sequentially followed by [ʂ] ($t = 2.63$, $p = 0.009$) and [ɕ] ($t = 8.62$, $p < 0.000$).

4 EMA results

The position data on the three sampled tongue points (tongue dorsum, tongue mid, and tongue tip) for the three SC sibilants were extracted at the time point when there is a tangential velocity minimum for the tongue tip. Figure 3 shows a midsagittal scatter plot of all repetitions of the data points for MS1 and FS2, respectively. The difference in shape denotes variations of the data points in clause position: (1) the downward triangle, circle, and cross for [s], [ʂ], and [ɕ] in the citation position, respectively; (2) the upper triangle, square, and plus for [s], [ʂ], and [ɕ] in the clause-mid position, respectively; (3) the left triangle, *, and pentagram for [s], [ʂ], and [ɕ] in the clause-initial position, respectively; (4) the right triangle, dot, and hexagram for [s], [ʂ], and [ɕ] in the clause-final position, respectively. And variations in vowel context are denoted by the grey scale: black for the sibilants in the [i] context, dark grey for the sibilants in the [a] context, and light grey for the sibilants in the [u] context, respectively.

As denoted by the superimposed 2-sigma confidence ellipses, [ɕ] is least variable, sequentially followed by [s] and [ʂ], in general. With reference to the speakers' palatal casts, it is found that the major variation axis is aligned parallel to the speakers' palate for [ɕ] whereas perpendicular for [s] and [ʂ], in general. As can be seen from the figure, there is no clear difference in clause position for the sibilants. The vowel context, however, may have an effect on lingual configurations for [s ʂ], especially as denoted by the two posterior tongue points: comparatively retracted for the [u] context and lowered for the [a] context vis-à-vis the [i] context. But for the tongue tip, there is no clear difference in vowel context. Also, the EMA data collaborate the EPG results in that (1) the tongue tip for [ʂ] is posterior to those for [s ɕ], (2) [s ɕ] exhibit no difference in terms of tongue tip, and (3) the three sibilants differ entirely in overall lingual configuration.

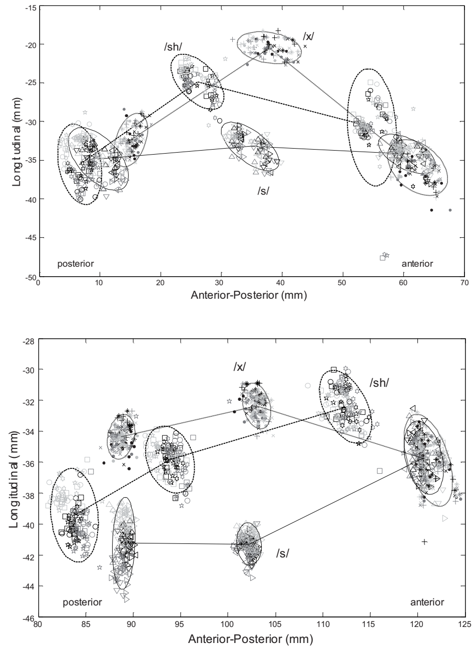


Figure 3: *Lingual configurations for /s/ ([s]), /sh/ ([ʂ]) and /x/ ([ɕ]) for MS1 (upper) and FS2 (lower).*

5 Acoustic results

Articulatory results have shown that there is no significant difference in clause position for the sibilants. That's probably attributable to the fact that the sibilants are all allocated in the stressed positions. However, acoustic data show that the clause position has an effect on sibilant duration in general. Figure 4 shows the box plots for the sibilant durations in different clause positions for MS1 (left) and FS2 (right), respectively. One-way ANOVAs yielded a significant effect of 'clause position' for MS1 ($F = 90.4$, $p < 0.000$) and for FS2 ($F = 95.49$, $p < 0.000$), respectively, except that there is no significant duration difference between the citation and clause-mid positions for FS2 ($t = -0.076$, $p = 0.94$). Interestingly, the two speakers exhibit a different ordering with respect to the sibilants duration: 'citation' > 'initial' > 'mid' > 'final' for MS1 whereas 'final' > 'citation' >= 'mid' > 'initial' for FS2. Meanwhile, within each speaker, all sibilants demonstrate a stable pattern of duration difference.

Figure 5 shows the box plots for the corresponding vowel durations. One-way ANOVAs yielded a significant effect of 'clause position' for MS1 ($F = 302.5$, $p < 0.000$) and for FS2 ($F = 373.1$, $p < 0.000$), respectively, except that there is no significant duration difference between the citation and clause-final positions for FS2 ($t = -0.406$, $p =$

0.685). For vowels, the two speakers exhibit a quite uniform ordering: 'citation' > 'final' > 'mid' > 'initial' for MS1 and 'citation' >= 'final' > 'mid' > 'initial' for FS2. That is, the final lengthening effect is generally attested for both speakers.

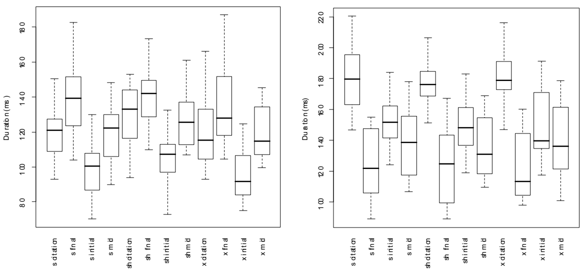


Figure 4: *Box plots of the SC sibilant durations for MS1 (left) and FS2 (right).*

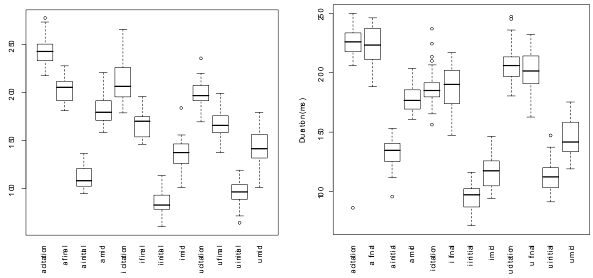


Figure 5: *Box plots of the corresponding vowel durations for MS1 (left) and FS2 (right).*

In summary, it seems that the prosodic effect on syllable durations mainly applies to the vowels, and the sibilants are more variable, especially between different speakers.

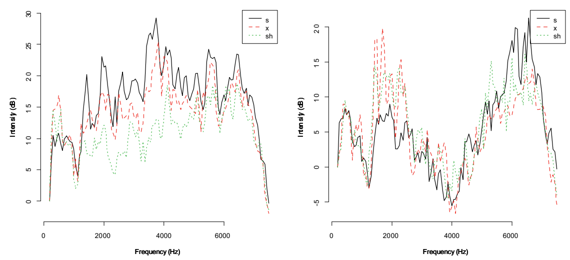


Figure 6: *Average spectra for the SC sibilants: MS1 (left); FS2 (right).*

A 256-point Hamming window was applied to the mid point on each sibilant interval (sampling rate = 16,000 Hz). Figure 6 gives the average FFT spectra for the three SC sibilants for MS1 (left) and FS2 (right), respectively. The two speakers exhibit different patterns, again. In MS1, [ʂ] demonstrates a flatter spectral envelope than [s ɕ], whereas [s ɕ] are not well distinguished from each other except that the latter has apparently more energy below 1 kHz than the former. Regarding FS2, (1) [ɕ] has the highest energy peak at around 1.8 kHz while a comparatively one at around 6.5 kHz, (2) [s] has the highest energy peak at around 6.5 kHz while a much lower one at around 1.8 kHz, and (3) the highest peak of [ʂ], interestingly, is patterned with that of [s], whereas its slightly lower peak at around 1.8 kHz is patterned with [ɕ].

6 Conclusion

(1) The three SC sibilants differ in overall lingual and linguopalatal configuration. (2) [s] and [ʂ] are clearly distinguished by place of articulation. (3) It's meaningless to characterize [ɕ] in terms of a distinct place of articulation. The IPA term of alveolo-palatal should thus be interpreted as the whole alveolo-palatal region, rather than somewhere in-between. (4) Speakers are the major source of articulatory as well as acoustic variability.

Acknowledgments

Work supported by CASS Key Project (Type A) "Aspects of Speech Production in Chinese" (No. 0500000265). And the author is grateful to Mr. Li Yunjing and Dr. Xiong Ziyu for their help in EPG data processing, and to Professor Eric Zee for making the palatograms and lingograms.

References

[1] Y. R. Chao. *Mandarin Primer*. Harvard University Press, Cambridge, MA. 1948.

[2] P. Ladefoged & Z. Wu. Places of articulation: an investigation of Pekingese fricatives and affricates. *Journal of Phonetics*, 12:267-278, 1984.

[3] W.-S. Lee. An articulatory and acoustical analysis of the syllable-initial sibilants and approximant in Beijing Mandarin. *Proc. ICPhS*, 413-416, 1999.

[4] I. Maddieson. *Patterns of Sounds*. Cambridge University Press, Cambridge, UK. 1984.

[5] M. Toda & K. Honda. An MRI-based cross-linguistic study of sibilant fricatives. *Proc. ISSP-6*, 2003.

[6] M. Toda. Deux stratégies articulatoires pour la réalisation du contraste acoustique des sibilantes /s/ et /ʃ/ en français. Paper resented at *the XXVIes Journées d'Étude sur la Parole (JEP)*, Dinard, France, 2006.

[7] S. Dart. Comparing French and English coronal consonant articulation. *Journal of Phonetics*, 26:71-94, 1998.

[8] J. Brunner, S. Fuchs & P. Perrier. The influence of the palate shape on articulatory token-to-token variability. *ZAS papers in Linguistics* 42: 43-66, 2005.

[9] D. Recasens & A. Espinosa. An electropalatographic and acoustic study of affricates and fricatives in two Catalan dialects. *Journal of the International Phonetic Association* 37(2):143-172.