SPECTRAL MEASURES FOR SIBILANT FRICATIVES OF ENGLISH, JAPANESE AND MANDARIN CHINESE

Fangfang Li[a], Jan Edwards[b], & Mary E. Beckman[a]
[a]Dept. Linguistics, Ohio State Uni. & [b]Dept. Communicative Disorders, Uni. Wisconsin -- Madison
{cathy | mbeckman}@ling.ohio-state.edu, jedwards2@wisc.edu

1.INTRODUCTION

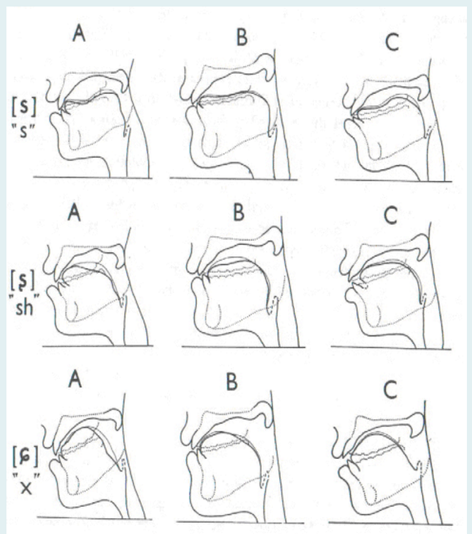
Table 1. The voiceless sibilants of English, Japanese and Mandarin

	Dental/ alveolar	Post-alveolar		
		Alveolo-palatal	Retroflex	Palato-alveolar
English	s			[ʃ]
Japanese	s	[ɕ]		
Mandarin	s	[ɕ]	[ʂ]	

actually involve **two different types of articulatory contrast** …

- **English /[ʃ]/ and /s/:** contrast in **constriction place**. The narrowest lingual constriction for /[ʃ]/ is further back than for /s/.
- **Japanese /[ɕ]/ and /s/:** contrast in **tongue posture** as well as in place. Not only is the constriction in /[ɕ]/ further back than the one in /s/, it also involves bunched predorsum, rather than the tongue tip/blade as in /s/.
- o **Mandarin /[ʂ]/, /[ɕ]/, and /s/:** contrast in **constriction place independently of tongue posture**. The postalveolar /[ʂ]/ and /[ɕ]/ contrast with /s/ in tongue position, and /[ʂ]/ contrasts with /[ɕ]/ and /s/ in tongue posture**.**

Figure 1. X-ray tracings of 3 Mandarin speakers producing /s-[ʂ]-[ɕ]/ (Ladefoged &Wu, 1984)



2. EXISTING ACOUSTIC MEASURES AND WHY THEY ARE NOT COMPLETELY SATISFATORY

2.1. Centroid frequency: A measure of constriction PLACE

- Used by Forrest et al. (1988), Jongman et al. (2000), Shadle & Mair (1996), among others.
- Defined as the first moment (mean frequency) when power spectrum is treated as a probability distribution.
- Interpreted as a measure of place, since it tends to be negatively correlated with the length of the front cavity in English sibilants.
- This interpretation fails for Mandarin /[ʂ]/, which has a slack constriction, allowing energy to leak from the back cavity (Fig. 3).

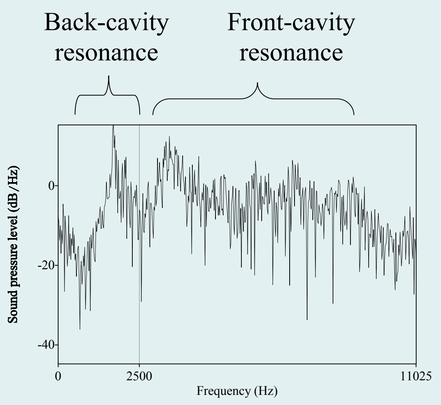


Figure 2. Spectrum of Mandarin /[ʂ]/ in *shí-zǐ* 'stone'. The spectral prominence below 2500 Hz reflects the long back cavity.

2.2 Onset F2 Frequency: A measure of constriction POSTURE

- Used by Funatsu (1995), Halle & Stevens (1997), among others.
- Defined as the F2 frequency taken at the onset of voicing in the vowel following the lingual fricative.
- Tends to be negatively correlated with the length of the back cavity.
- Exact value depends on following vowel, making cross-language comparison difficult when vowel systems differ.
- Value not always available – e.g., Tokyo Japanese vowels are frequently deleted or devoiced after voiceless fricatives.

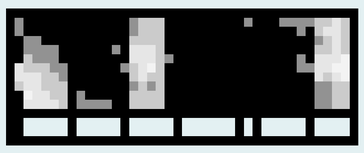


Figure 3. Production of *sakana* 'fish'by a male adult speaker of Tokyo Japanese, with the first [a] devoiced.

3. MATERIALS AND DATA COLLECTION

- **Materials:** Audio recordings of words beginning with target fricative followed by vowels /a/, /i/, /u/, /e/, /o/.
- **Participants:**10 speakers for each language (5 males and 5 females), recorded in Tokyo, Japan, Columbus, US, and Songyuan, China.

4. DEVELOPING NEW ACOUSTIC MEASURES

- Based on separating contribution of front cavity resonance in higher frequencies from contribution of back cavity in the F2 region (see Fig. 4).
- F2 region: Ideally, this is the F2 in the fricative or in [[ʂ]]. Since there is no observable F2 peak in [s] or [[ʂ]], and no true neutral vowel in these three languages, the F2 region was approximated by taking the F3 of /a/ and multiplying by 3/5 for each speaker in each language.

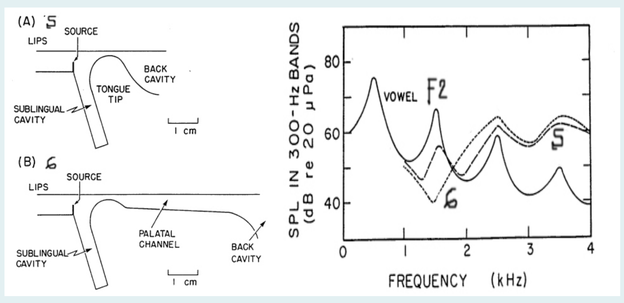
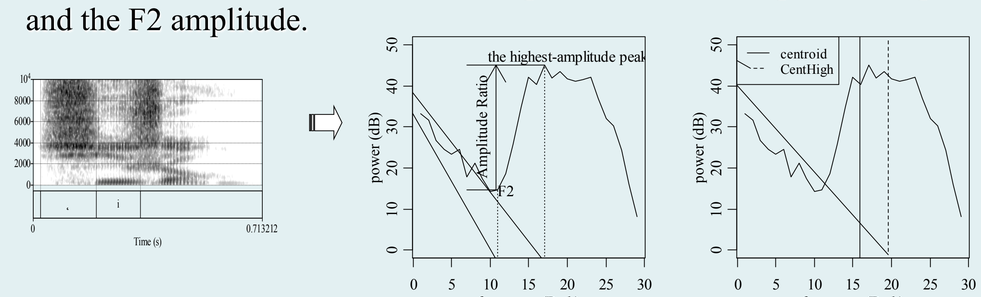


Figure 4. Cavity shapes for /[ʃ]/ and /[ɕ]/ (left panel), with /[ʃ]/ having a short slack constriction, whereas /[ɕ]/ involves a long palatal channel instead. The right panel is the theoretical spectra of /[ʃ]/ and /[ɕ]/ overlaid on the spectrum of a neutral vowel (from Halle & Stevens 1997).

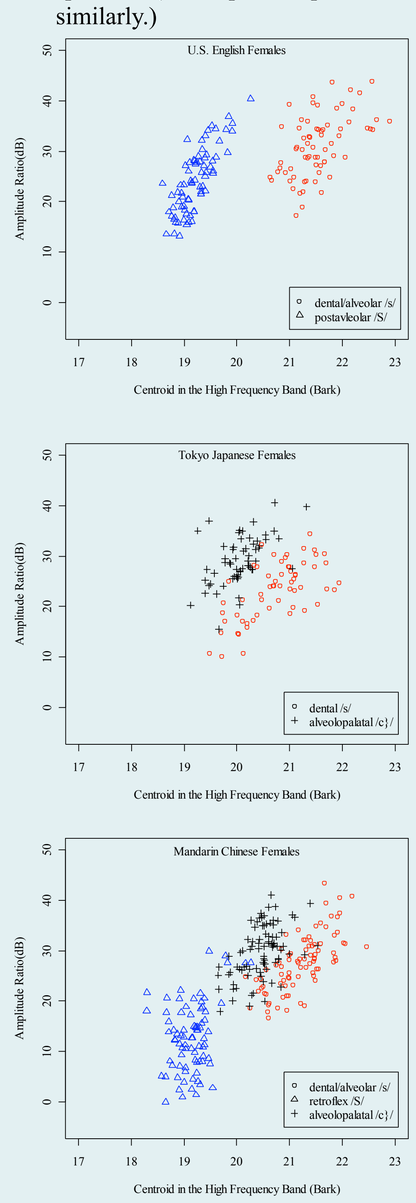
- **Centroid in the Higher Frequency Band (CentHigh):** A measure of constriction **PLACE**. Refinement of the centroid frequency measure so that it is now calculated over a higher frequency band above "F2 region".
- **Amplitude Ratio (ampRatio):** A measure of tongue **POSTURE.** It is the difference in dB between the amplitude of the most prominent peak and the F2 amplitude.



5. RESULTS

- **English:** /s/ and /[ʃ]/ differentiated by centHigh, with /s/ having higher values than /[ʃ]/ (Fig. 5, top).
- **Japanese:** /s/ and /[ɕ]/ differentiated by both centHigh and ampRatio (Fig. 5, middle). Linear discriminant analysis showed both parameter to correctly identify 92% of the tokens.
- **Mandarin:** The place distinction was captured by centHigh, and the posture difference was captured by ampRatio (Fig. 5, bottom). Linear discriminant analysis again correctly identified 92% of the tokens.
- The typically apico-alveolar /s/ of English had higher centHigh and higher ampRatio than the lamino-dental /s/ of Japanese and Mandarin.

Figure 5. AmpRatio plotted against CentHigh for English, Japanese and Mandarin female speakers. (Male speakers pattern similarly.)



6. DISCUSSION

- CentHigh and ampRatio are two parameters that can effectively distinguish two types of articulatory contrasts: contrast in tongue position and contrast in tongue posture.
- The /s/ phonemes in the three languages differ in phonetic detail, in a way that maximizes the contrast with the alveolo-palatal fricative in the inventory.

Acknowledgement

This work was supported by NIDCD grant 02932 to Jan Edwards. We thank the speakers who participated in this study. We also thank Laura Slocum and Junko Davis, Kiyoko Yoneyama, and Yuki Sunawatari, who collected the English and Japanese data.