ACOUSTIC AND AUDITORY COMPARISONS OF POLISH AND TAIWANESE MANDARIN SIBILANTS

Chenhao Chiu
Dept. of Linguistics, University of British Columbia, 2613 West Mall, Vancouver, BC, Canada V6T 1Z4

1. INTRODUCTION

Polish and Mandarin Chinese are two languages well known for the three-way distinction of sibilant fricatives [1]: apical dental [s], laminal flat post-alveolar (retroflex[a]) [⍰], and laminal palatalized post-alveolar (alveolo-palatal) [⍰]. The spectral measurements of English, Japanese and Mandarin sibilants in Li, Edwards, and Beckman [2] show that sounds with the same IPA transcription can be phonetically different and the differences are dependent on what contrasts are present in the languages inventories. McGuire (2007) [3] points out that Mandarin speakers are better than American English speakers at using frication noise and formant transition as cues to distinguish Polish retroflex and alveolo-palatal sounds. The present study compares the acoustic properties of the three sibilants in Polish (PL) and Taiwanese Mandarin (TM) and discusses the differences in terms of frication noise and the formant transitions. Aside from the auditory and articulatory similarities described in Ladefoged and Maddieson [1], this study will try to answer two main questions: (1) Are the sibilant contrasts in PL and TM the same? If not, how are they acoustically different? (2) What acoustic cues and properties distinguish one sibilant from another in these two languages? This study provides acoustic data, including moments, frequencies, formant transitions and LPC spectra showing that there exist the contrasts of the sibilant types in the two languages.

2. METHOD

2.1 Materials

The dental sibilant and the alveolo-palatal are in complementary distribution in both languages. The alveolo-palatal [⍰] only occurs before a high front vowel ([i] for PL; [i] and [y] for TM) or a glide ([j] for PL; [j] and [ɥ] for TM) whereas the dental sibilant [s] occurs elsewhere. Both PL and TM retroflex [⍰] occur in the same environment as the dental sibilant [s]. With these phonological conditions considered, the material environments for [s] and [⍰] and for [⍰] are designed to be different. In this study, the materials are real words[b] starting with a sequence of sibilant + vowel. [s] and [⍰] are followed by a low back vowel [a] (PL: [sab⍰t] "sabot", [⍰abla] "saber"; TM: [sa] "to sprinkle", [⍰a] "to kill"). [⍰] precedes a glide [j] followed by a low back vowel [a] (PL: [⍰jada⍰] "to sit"; TM: [⍰ja] "blind"). Fifteen tokens of each sibilant type were recorded by one male PL native speaker and two male TM native speakers.

[a] The present study follows Ladefoged and Maddieson (1996) in assuming Polish and Mandarin retroflexes to be laminal post-alveolar.

[b] Since TM does not have coda consonants (except for the alveolar nasals and velar nasals), only monosyllabic words are recorded.

2.2 Analysis preparation

For each token, the analyses included the centre of gravity (CoG) and F2 transitions of the sibilants along with the dynamic amplitude and frequencies of the maximum amplitude, as applied in Jesus and Shadle [4]. A Hanning window of 100ms from the midpoint of each token was selected to produce LPC spectra. The recording sampling rate was 44.1 kHz. The analysis window length was 25ms. The second formant (F2) of the following vowel ws measured in the beginning 25ms and the middle 25ms of the vowel (Hanning windowed). The measurements of the dynamic amplitude and the frequency of the maximum amplitude follow Jesus and Shadle [4]. The dynamic amplitude is the difference between the maximum amplitude between 500 Hz – 20K Hz and the minimum amplitude between 0 and the maximum amplitude. The frequencies of the maximum amplitude were also measured. All the measurements were performed with the phonetic software Praat.

3. RESULTS

3.1 Acoustic analyses[c]

Table 1 gives the mean CoG values and the frequencies of the maximum amplitudes of sibilants. For both languages, a main effect of sibilant types is found in CoG (PL: $F(2, 42) = 299.645$, $p < .001$; TM: $F(2, 42) = 54.056$, $p < .001$). Pair-wise t-tests within places of articulation also reach significance (for [s] pair: $p = .007$; [⍰]: $p < .001$; [⍰]: $p < .001$). For frequencies of the maximum amplitudes, substantial differences are found in the retroflex pair and the alveolo-palatal pair (both $p < .001$). For both the CoG and maximum frequencies, within-language analyses show that in PL while dentals are different from retroflexes and alveolo-palatals, retroflexes and alveolo-palatals are not significantly different whereas all TM sibilants are significantly different.

Table 1. Mean CoG and frequency of the maximum amplitude

Context	Center of gravity (Hz)		Max Freq. (Hz)	
	PL	TM	PL	TM
[s]	8442.07	7734.15	8354.03	8079.8
[⍰]	2287.03	5598.7	2811.74	4597.38
[⍰]	2788.7	6288.55	3391.62	6098.49

When comparing F2 transitions (as shown in Table 2), a significant difference between PL and TM is found only in the retroflex pair ($p < .001$). Post hoc analyses found

[c] All the moment values and frequencies from the two TM speakers are highly correlated, except for the F2 transitions. In present study, the average values of the two speakers are used for analyses.

substantial differences between PL dentals and alveolo-palatals (p <.001) as well as among all three sibilants in TM (all p <.001). The dynamic amplitudes of the dental pair and alveolo-palatal pair between the two languages are significantly different (both p <.001), but not the retroflex pair (p = 0.121). In addition, main effects of sibilant types are also found in both languages (PL: p =.002; TM: p <.001). Post hoc tests show that in terms of dynamic amplitude, PL [s] and TM [⍰] are significantly different from their counterparts.

Table 2. Mean F2 transition and dynamic amplitude

Context	F2 transition (Hz)		Dynamic amplitude (dB)	
	PL	TM	PL	TM
[s]	58.93	64.18	38.86	29.93
[⍰]	71.12	170.05	26.65	32.2
[⍰]	357.44	337	28.47	38.08

The author observed lip protrusion in PL [⍰] and [⍰]. For TM, lip protrusion is optional for [⍰] and is obligatory for [⍰] only when it precedes [y]. Another fifteen tokens of TM [⍰y] were recorded for comparison (CoG: 5756.67 Hz; Max. Freq.: 5010.21 Hz; dynamic amplitude: 35.8 dB). The acoustic properties of PL [⍰] and TM rounded [⍰] are significantly different (CoG: p <.001; Max. Freq.: p =.001; dynamic amplitude: p =.002).

3.2 Spectra comparison

Among all the sibilant types, the retroflex pair in the two languages is the most distinctive. Their CoG, maximum frequencies, F2 transitions, and dynamic amplitude are significantly different. Nevertheless, the LPC spectra show that they are acoustically different (Fig. 1), but auditorily similar (Fig. 2).

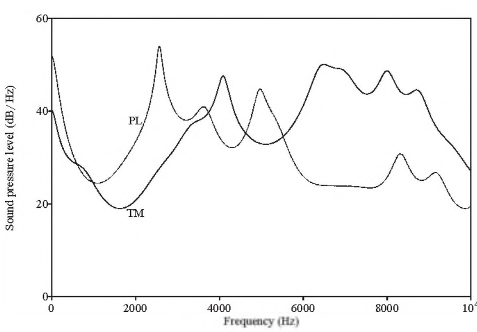

Figure. 1. Acoustic LPC Spectrum of PL [⍰] and TM [⍰]

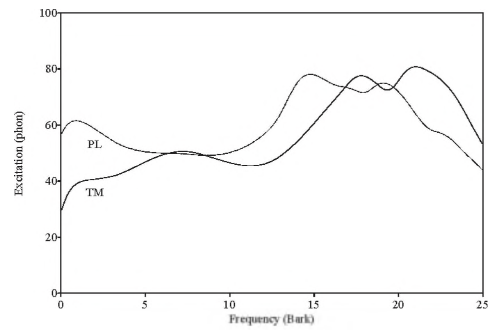

Figure 2. Auditory LPC Spectrum of PL [⍰] and TM [⍰]

4. DISCUSSION

Despite the fact that dental sibilants in the two languages are quite similar with respect to CoG and frequency of the maximum amplitude, the acoustic properties are rather different in PL and TM sibilant contrasts, even though the same IPA transcriptions are used for both languages. For both languages, the CoG and frequencies of the maximum amplitude are used to distinguish the anterior sibilant from the other two types and their cross-language differences are yet substantial. The lip protrusion contributes to the F2 lowering and the defining characteristic of stronger spectral energy at lower frequency ranges for PL [⍰] and [⍰], distinct from TM rounded [⍰] and unrounded [⍰]. These gestural variations seem influential to the acoustics of sibilants. Ladefoged and Maddieson [1] point out the auditory similarities of sibilants in these two languages. However, the acoustic effect of F2 transition stronger in TM than in PL. Although frication noise and the formant transitions are the most salient acoustic contrasts in PL (consistent with the findings in [5]) and TM, whether or not these acoustic cues are determinative to the perception requires more empirical evidence.

REFERENCES

[1] Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Blackwell. Oxford.

[2] Li, F., Edwards, J., and Beckman, M. (2007). Spectral measures for sibilant fricatives of English, Japanese and Mandarin Chinese. *Proceedings of the XVIth International Congress of Phonetic Sciences*, pp. 917 – 920.

[3] McGuire, G. (2007). Phonetic cue and integrality. *JASA*. 122, 5, Nov., 2971.

[4] Jesus, L. M. T. and Shadle, C. H. (2002). A parametric study of the spectral characteristics of European Portuguese fricatives. *Journal of Phonetics*, 30, 437 – 464.

[5] Padgett, J. and Zygis, M. (2006). *A Perceptual Study of Polish Fricatives, and its Relations to Historical Sound Change*. Poster presented at 10th Conference on Laboratory Phonology. Paris, 29.06 – 01.07.2006.

ACKNOWLEDGEMENTS

Special thanks to Amanda Miller and Molly Babel for comments and suggestions, and to Carrie Cuttler and Takako Kawasaki for their assistance on statistic analyses.