The acoustic analysis of Mandarin Sibilants in the Production of Hungarian learners of Chinese

Kornélia Juhász

Eötvös Loránd University, Hungarian Research Centre for Linguistics, Hungary
juhasz.kornelia8@gmail.com

Abstract

In this study we carried out an acoustic analysis on the three voiceless Chinese sibilants, that is the dental [s], the alveolo-palatal [ɕ] and the so-called retroflex post-alveolar [ʂ], by comparing the production of Hungarian learners of Mandarin Chinese and Chinese natives. We analysed the three Chinese sibilants according to two parameters: the sibilant constriction's relative position in the vocal tract by the length of the cavity anterior to the sibilant's constriction, and the degree of palatalization, that is, the length of the cavity posterior to the sibilant constriction. Due to the distinct distribution and opposition of sibilants in Hungarian and Chinese, we assumed negative transfer-induced difficulties in Hungarian speakers' production in the case of the alveolo-palatal [ɕ]. Our results confirmed this hypothesis that Hungarian learners produced the [ɕ] identical to the dental [s] and the (alveolo-)palatalized production (which is absent from the L1) is substituted with a coarticulatory-driven gradually palatalizing pattern at the end of the sibilant phase. The production of the dental [s] and the so-called retroflex post-alveolar [ʂ] did not pose problems for Hungarian learners of Chinese.

Keywords: acoustic analysis, L2 speech sound production, Chinese sibilants

1. Introduction

This study focuses on the acoustic analysis of the Mandarin Chinese sibilants' production among Hungarian learners of Chinese. In terms of sibilant contrast, Chinese and Hungarian differ: in Chinese there is a three-way place contrast, as there are the dental [s], the alveolo-palatal [ɕ] and the so-called retroflex post-alveolar [ʂ] [1, 2]. The so-called retroflex post-alveolar [ʂ] features an apical or predorsal production accompanied with a retracted tongue root and a sublingual cavity [1, 2]. However, in Hungarian there is only a two-way contrast between a dentialveolar [s] and a post-alveolar [ʃ] [3]. Due to the different distribution and opposition of the two languages' sibilants, we assume that the Chinese sibilant contrast poses problems to Hungarian speakers [4]. To test this assumption, we carried out an acoustic analysis on the above mentioned three segments comparing the production of Hungarian learners of Chinese and Chinese native speakers. We analysed the three segments according to two parameters. Firstly, we analysed the place of articulation, that is to say, the sibilant constriction's relative position in the vocal tract by the length of the cavity anterior to the sibilant's constriction. Secondly, we analysed the degree of palatalization, that is, the length of the cavity posterior to the sibilant constriction.

In the production of sibilant speech sounds, acoustic structure is fundamentally defined by the cavity anterior to the constriction, due to the fact that this cavity's length is responsible for the acoustic quality of the sibilant's turbulent noise [5]. In the acoustic analysis of sibilants, centre of gravity (COG) can be correlated negatively with the length of the cavity anterior to the sibilant constriction [1]. Among Chinese sibilants, the alveolo-palatal [ɕ] possesses an intermediate COG value relative to the dental [s] and the retroflex post-alveolar [ʂ] (Figure 1.). More particularly, in the production of Chinese speakers, [ɕ] had lower COG values, than the dental [s], but higher COG than the post-alveolar [ʂ] [1, 6].

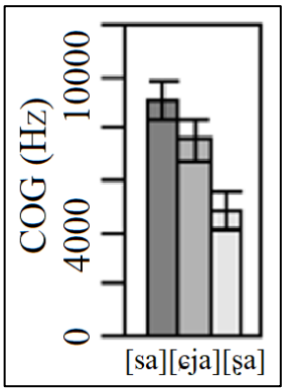

Figure 1. The Centre of Gravity (Hz, measured at the middle of the fricative phase) of the Chinese dental [s], the alveolo-palatal [ɕ] and the so-called retroflex post-alveolar [ʂ] ([1]: 42)

10.21437/ISAPh.2021-8

It is worth to note that the intermediate value of [ɕ] can be accounted for by the position of the constriction, which is according to [7]'s palatogram study, placed at an intermediate position between the dental [s] and the post-alveolar [ʂ]. However, we should also point out that, apart from the front cavity's resonance, we cannot exclude the possibility of the involvement of the narrow palatal channel in the production of the alveolo-palatal [ɕ], as it was observed in the case of the Polish counterpart of the alveolo-palatal [ɕ] [5].

In terms of COG manipulation we can differentiate between two distinct patterns of palatalization [8]. There is the already mentioned Chinese (alveolo-)palatalized pattern, where the [ɕ] carries intermediate COG throughout the whole fricative phase. However, there is another, coarticulation-induced palatalizing pattern (e.g. in English), which features COG frequency decrease only in the second half of the fricative interval (compared to the realization of a sibilant placed to a non-palatal context) as the constriction is moving backwards in the oral cavity, and lengthens the front cavity [8] (Figure 2).

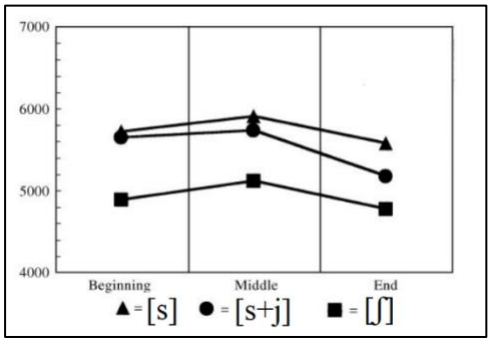

Figure 2. Coarticulatory palatalizing COG-pattern (HZ) of [s] in English, compared to realization of the [s] and [ʃ] with non-palatal context ([8]: 85)

Aside from COG, the degree of palatalization can also be measured from a different point of view, at the transition between the sibilant and the following vocalic section using the initial second formant (F_2). In the case of coronal segments, the initial F_2 is negatively correlated with the length of the back cavity [1, 9]. In the production of the alveolo-palatal [ɕ] sibilant, the tongue is bunched and a long and narrow channel is created, which implies a strongly palatalized production [1, 2]. Since the alveolo-palatal [ɕ] features the smallest back cavity among Chinese sibilants, it possesses the highest F_2 relative to dental [s] and the retroflex [ʂ]. Due to the relatively larger length of the posterior cavity, initial F_2 is lower both in the dental [s] and the retroflex [ʂ] than in the alveolo-palatal [ɕ] [1] (Figure 3.). To be concrete, the initial F_2 is the lowest in the case of the dental [s], because the constriction is placed at a more anterior position, while the post-alveolar [ʂ] has intermediate F_2, as the constriction is more backwards in the oral cavity, but the tongue is not bunched to the palate.

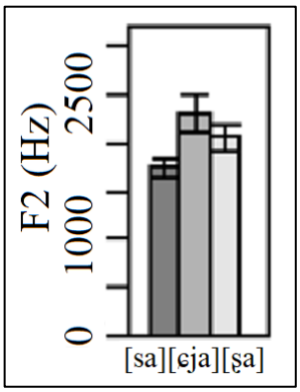

Figure 3. The initial F_2 (Hz, measured in a 25 ms window beginning from the appearance of the full formant structure) of the Chinese dental [s], so-called retroflex post-alveolar [ʂ], and the alveolo-palatal [ɕ] ([1]: 42)

2. Hypotheses

The Chinese dental [s] and the retroflex post-alveolar [ʂ] have a counterpart in Hungarian, namely the dentialveolar [s] and the post-alveolar [ʃ]. According to phonetic descriptions of these two sound pairs, they have identical places of articulation in the two languages, respectively. [1, 3]. On this basis, we hypothesize that Hungarian learners of Chinese produce these two Chinese segments the same way as Chinese native speakers, that is, COG values measured in these segments do not differ in Hungarian and Chinese speakers. Additionally, we do not expect these two sibilants produced by Hungarian speakers to differ from those of Chinese speakers in terms of initial F_2 values either, because we do not expect any differences in the length of the back cavity, i.e. in the degree of palatalization.

As Hungarian lacks the segment [ɕ], but also the articulatory feature of palatalization [3], we assume that there is a negative transfer effect, more particularly a single-category assimilation [4]. On this basis, we expect Hungarian learners of Chinese to assimilate the alveolo-palatal [ɕ] to one Hungarian speech sound category, namely to the Hungarian dentialveolar [s]. The assimilation in this case is attributed to the effect of orthography [10]: in pinyin (the romanization of Chinese characters), the alveolo-palatal [ɕ] is denoted by a grapheme '*x*' which, in Hungarian orthography, denotes the [ks] sequence containing the speech sound [s]. Accordingly, we expect Hungarians to produce the Chinese alveolo-palatal [ɕ] similar to how they produce the Chinese dental [s], however we suppose a natural coarticulatory palatalizing process, induced by the

following palatal vocalic segment, which is expected to cause a significant decrease in frequency at the second half of the sibilant phase (as opposed to the dental [s], where the palatal context is absent, thus there is no effect of palatalization) [8, 11, 12]. Due to the fact that this way the palatalizing process is carried out in the sibilant phase, we expect Hungarian speakers to produce [ɕ] with the same initial F_2 as that of Chinese native speakers.

3. Method

We compared the production of two groups: one group of Hungarian learners of Chinese (2nd year university students majoring in Chinese, learned approximately 1.5 year of Chinese), and Chinese native speakers (5 speakers per group, women). We analysed the dental [s], the alveolo-palatal [ɕ], and the retroflex post-alveolar [ʂ] in isolated one-syllable (C(G)V structured) Chinese words that were read three times in a randomized order by all speakers. The three Chinese sibilants appeared in four different phoneme context, /a/, /u/, /i/, /ɤ/, (4 contexts * 3 repetitions * 10 speakers * 3 sibilants, altogether 360 samples) and were projected with Chinese characters and pinyin as well (Table 1.).

Table 1: The analysed Chinese words with characters, pinyin, and their meaning

	Dental [s]	**Alveolo-palatal [ɕ]**	**Retroflex [ʂ]**
/i/	四 sì [sɨ̂] 'four'	系 xì [ɕî] 'department'	是 shì [ʂɨ̀] copula
/a/	萨 sà [sâ] 'a surname'	夏 xià [ɕjâ] 'summer'	厦 shà [ʂâ] 'mansion'
/u/	速 sù [sû] 'speed'	绣 xiù [ɕjû] 'excellent'	数 shù [ʂû] 'number'
/ɤ/	色 sè [sɤ̂] 'color'	谢 xiè [ɕjɛ̂] 'to thank'	射 shè [ʂɤ̂] 'to shoot'

The acoustic analysis was carried out by the altered version of [13]'s script. We measured COG at three different timepoints throughout the sibilant, at 20%, 50%, and 80% of the segment's total duration, which enabled us to reveal any gradual changes that might occur in the constriction's relative position within the fricative interval [8]. F_2 was measured in a 25 ms window starting from the appearance of the full formant structure, at the beginning of the vocalic segment [1]. The acoustic analysis was carried out in Praat [14], data were submitted to statistical analysis, linear mixed effects models were used [15] in R [16]. The dependent variable of COG was analysed by the interaction effect of independent variables of speaker group (Hungarian/Chinese), place of articulation (dental/alveolo-palatal/retroflex) and time point (20%/50%/80%). Speaker was defined as random variable. The pairwise comparison was carried out with post hoc tests [17]. In the case of the dependent variable initial F_2, the same model was applied, however without the independent variable of timepoint, due to the fact that only a single value was measured here. The figures were visualized by the ggplot2 package [18].

4. Results

In terms of COG the linear mixed model showed a significant interaction effect among the three independent variables, that is, speaker group, place of articulation and timepoint ($F(4,1067) = 2.86$, $p < 0.05$). As it was expected, the dental [s] had the highest, while the post-alveolar [ʂ] had the lowest COG means among the three Chinese sibilants (Figure 4.). The Chinese speakers discriminated all Chinese segments significantly in terms of COG, at all timepoints measured ($p < 0.001$). However, Hungarian speakers produced the alveolo-palatal [ɕ] identically to the dental [s] at the first two timepoints measured, that is, at the 20% and 50% of the duration of the sibilant phase. Hungarian speakers only differentiated the alveolo-palatal [ɕ] and the dental [s] at the 80% timepoint, where the [ɕ] had significantly lover COG than the dental [s] ($p < 0.001$). We found significant differentiation of all other sibilants and at all timepoints measured in both groups.

Comparing Hungarians' and Chinese natives' sibilant production, the only difference was found at the 50% timepoint of the alveolo-palatal [ɕ], because Hungarians produced this segment with a significantly higher COG than Chinese natives ($p < 0.001$). We did not find any other significant difference in all other sibilants and timepoints compared. Concerning tendencies, although no significant difference was found, however as a tendency Hungarians produced the so-called retroflex post-alveolar [ʂ] with lower COG means than Chinese natives.

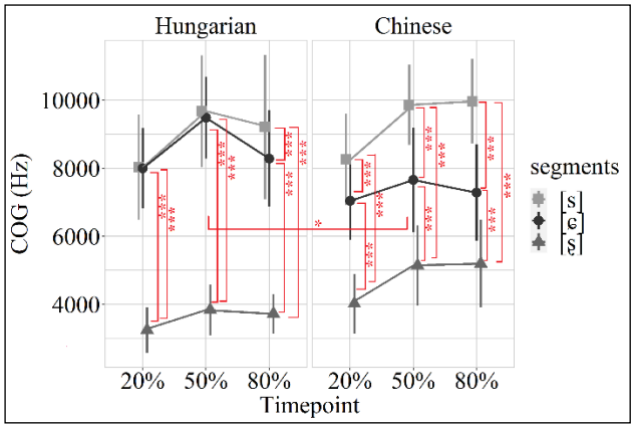

Figure 4. The Chinese dental [s], alveolo-palatal [ɕ] and retroflex post-alveolar [ʂ] segments' COG values (means and SDs) in the production of Hungarian and Chinese speakers

Regarding the dependent variable F_2, the mixed model showed a significant interaction effect between the two independent variables, that is, the speaker

group and place of articulation ($F(2, 350) = 15.65$, $p < 0.001$). In both speaker groups, the dental [s] had the lowest, and the alveolo-palatal [ɕ] had the highest, while the so-called retroflex post-alveolar [ʂ] was realized with intermediate F_2 means (Figure 5.).

Regarding the within-group comparison, both speaker groups discriminated the alveolo-palatal [ɕ] from the other two Chinese sibilants ($p < 0.001$). However, while Chinese natives produced the dental [s] and so-called retroflex post-alveolar [ʂ] distinctly, Hungarian speakers did not produce such differentiation, thus Hungarians realized these two segments identically in terms of F_2 measured at the beginning of the vocalic segment. Comparing F_2 of the three sibilants across groups, that is, between Hungarians and Chinese natives, we found no significant differences in the case of either of the sibilants.

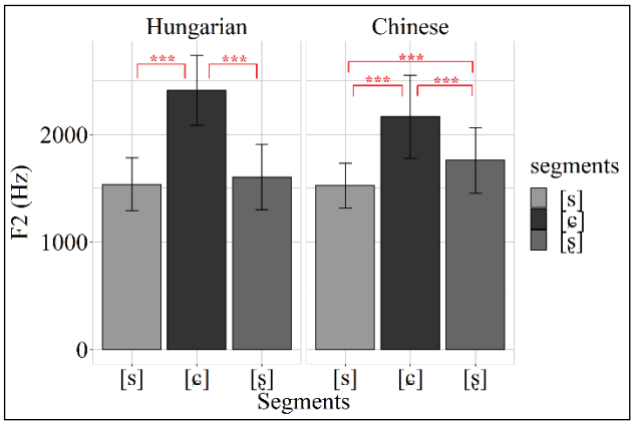

Figure 5. The Chinese dental [s], alveolo-palatal [ɕ] and retroflex post-alveolar [ʂ] segments' initial F_2 values (means and SDs) in the production of Hungarian and Chinese speakers

5. Discussion

In this study we analysed three Chinese sibilants, the dental [s], the so-called retroflex post-alveolar [ʂ] and the alveolo-palatal [ɕ] in the production of Hungarian learners of Chinese. We analysed two different acoustic parameters of these Chinese sibilants: the length of the cavity anterior to the sibilant's constriction by measuring COG at three different timepoints within the sibilant phase, and we also measured the degree of palatalization with the initial F_2, which is correlated with the length of the posterior cavity relative to the fricative constriction. We compared the production of a Hungarian group with a Chinese native control group. The target sibilants were read within one syllable Chinese words with four different vowel contexts.

We hypothesised that due to the identical places of articulations of the dental [s] and post-alveolar [ʂ] segments in Hungarian and Chinese, we did not expect differences between the Hungarian language learners' and the Chinese natives' production. This hypothesis is confirmed by our results, because we did not find any significant difference between the two groups either in terms of COG, or in initial F_2. It is worth to mention however that Hungarians produced lower COG as a tendency at each time point in the production of the so-called retroflex post-alveolar [ʂ] (compared to natives). This observed difference could result from distinct labial features in the two languages, that is, if we assume the transfer of a more labial Hungarian pattern of the Hungarian post-alveolar [ʃ], this way lengthening the vocal tract could cause lower COG, as the constriction's position is realized relatively more backwards in the oral cavity [19]. This feature could also give explanation to the lower initial F_2 as a tendency, because, due to this coarticulatory effect, the relatively more labial production lowers the formant frequencies; in this case, it particularly lowers the F_2 of the following segment as well [20]. This more labial production might be also the reason for the lack of differentiation between the dental [s] and the so-called retroflex post-alveolar [ʂ] in Hungarian learners' production.

Regarding the alveolo-palatal [ɕ], our results confirmed the effect of negative transfer, that Hungarians produced the alveolo-palatal [ɕ] identically to the dental [s] in the first half of the sibilant phase, and only differentiated the two segments, [s] and [ɕ] at the end of the turbulent noise. This differentiation can be accounted for the different contexts in the dental [s] and the alveolo-palatal [ɕ] occur. Due to its palatalized production, the alveolo-palatal sibilant is always followed by a palatal segment, however the dental [s] is followed by anything else but a palatal segment, as they are in complementary distribution [11, 12]. In accordance with these, in Hungarian speakers' production, the expected coarticulation-induced palatalizing COG-lowering occurred in the latter half of the alveolo-palatal [ɕ], however in the case of the dental [s], such lowering effect was not observed, owing to the lack of palatal context and thus palatalizing effect. In this clear case of single-category assimilation, Hungarians produced the alveolo-palatal [ɕ] visibly identical to the dental [s], thus the secondary (alveolo)-palatal feature was absent. As a result we have no reason to consider further differences between groups regarding the distinct involvement of the narrow palatal channel throughout the production. Due to the fact that the palatalizing process was carried out completely in the sibilant phase, thus the production of the vocalic segments was not affected by the distinct sibilant realizations in the two groups. Hungarian speakers' higher F_2 as a tendency may be attributed to a more bunched tongue dorsum

compared to Chinese natives, which occupies more space from the back cavity.

6. Conclusions

The production of those Chinese sibilants that feature identical places of articulation respectively in L1 and L2, that is, the dental [s] and the so-called retroflex post-alveolar [ʂ] did not pose problems for Hungarian L2 learners. Our negative transfer hypothesis regarding the Hungarian speakers' production of the Chinese alveolo-palatal [ɕ] is confirmed by our results and shed light to some aspects of the L2 speech sound acquisition process. Palatalization, as a secondary articulatory gesture is absent from Hungarian, hence, in order to resolve this shortcoming, language learners replace this secondary feature by a natural coarticulatory palatalizing pattern, only affecting the end of the sibilant phase. In sum, although the palatalization was not identical to the Chinese pattern, it was carried out, similarly to the natives' production, completely within the sibilant phase. Thus it did not affect the production of the following vocalic segment in either group. Accordingly, we can conclude, that the single-category assimilation is confirmed, the distinct distribution and opposition of L1 and L2 sibilants does affect the production of language learners, however coarticulatory modulations were carried out in order to align and not violate the transition to the following vocalic segment. The results of this study contribute to the understanding of the phenomenon of language transfer and target language speech sound acquisition.

7. Acknowledgements

I am grateful to my supervisor, Andrea Deme and Alexandra Markó for their kind suggestions and support.

This study was supported by the ÚNKP-21-3 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

8. References

[1] Li, F. 2008. *The phonetic development of voiceless sibilant sibilants in English, Japanese and Mandarin Chinese*. PhD Dissertation. Ohio: Ohio State University.

[2] Ladefoged, P., Wu, Z. J. 1984. Places of articulation: an investigation of Pekingese sibilants and affricates. *Journal of Phonetics.* 12(3), 267-278.

[3] Bolla, K. 1995. Magyar Fonetikai Atlasz. Budapest: Nemzeti Tankönyvkiadó Rt.

[4] Best C. T. 1995. A direct realist perspective on cross-language speech perception. In: Strange, W. (ed), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research.* York: Timonium, MD, 167-200.

[5] Toda, M., Maeda, S., Honda, K. 2010. Formant-cavity affiliation in sibilant sibilants". In: Fuchs, S., Toda, M., Zygis, M. (eds), *Turbulent Sounds: An Interdisciplinary Guide*. Berlin: De Gruyter Mouton. 343-375.

[6] Lee-Kim S. I. 2011. Spectral Analysis of Mandarin Chinese Sibilant Fricatives. *ICPhS XVII.* 1178-1181.

[7] Ladefoged, P., Maddieson, I. 1996. *The sounds of the world's languages*. Oxford: Blackwell.

[8] Zsiga, E. C. 2000. Phonetic alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28, 69-102.

[9] Nittrouer, S., Studdert-Kennedy, M., McGowan, R. 1989. The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research* 32. 120-132.

[10] Bassetti, B. 2008. Orthographic input and second language phonology. In: Piske, T. and Young-Scholten, M. (eds.), *Input Matters in SLA.* Clevedon, UK: Multilingual Matters 191-206.

[11] Lin, Y. H. 2007. *The Sounds of Chinese*. Cambridge: Cambridge University Press.

[12] Duanmu, S. 2000. *The phonology of standard Chinese*. New York: Oxford University Press.

[13] Styler, W. 2014. Spectral COG Finder Script. https://github.com/stylerw/styler_praat_scripts/blob/master/peak_spectral_COG.praat (downloaded: March 13th 2021).

[14] Boersma, P., Weenink, D. 2019. Praat: doing phonetics by computer [Computer program]. Version 6.1.05. (downloaded: November 4th 2019).

[15] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effect Models. *Journal of Statistical Software* 82(13), 1-26.

[16] R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, https://www.R-project.org/. 3.6.1-es verzió. 2019. (downloaded: November 4th 2019).

[17] Lenth, R. 2020. Emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.4. Elérhető: https://CRAN.R-project.org/package=emmeans. (downloaded: March 13th 2021).

[18] Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

[19] Harrington, J., Cassidy S. 1999. *Techniques in Speech Acoustics.* Dordrecht: Kluwer Academic Publishers.

[20] Stevens, K. N. 2000. *Acoustic Phonetics*. Massachusetts: MIT Press.