THE PERCEPTION OF MANDARIN SIBILANTS BY JAPANESE SPEAKERS: PREDICTION BY PAM

Yongzhe Peng

Sophia University
tetsu.peng.2015@gmail.com

ABSTRACT

As a theoretical model for cross-language perception, the Perceptual Assimilation Model accounts for listeners' discrimination of non-native contrasts from an unknown language. For Japanese listeners, whose native language has a two-way sibilant contrast, the perception of Mandarin sibilants, with a three-way contrast, is of interest, although not fully explored so far.

Data from ten Japanese speakers without any experience of Mandarin learning were collected for two experiments. In experiment I, the listeners were asked to select the most similar Japanese sound with the target stimuli and to rate its goodness on a 5-point scale. In experiment II, an AXB discrimination task of the three Mandarin sibilant contrasts was conducted. The results indicated that the discrimination of Mandarin retroflex /ʂ/ and alveolo-palatal /ɕ/, whose assimilation pattern is Single Category assimilation, is relatively difficult for Japanese listeners with no Mandarin experience, consistent with PAM's prediction.

Keywords: Mandarin, sibilant, naïve listeners, sound discrimination, perceptual assimilation

1. INTRODUCTION

The present study investigates the perception of Mandarin three-way sibilant contrasts by Japanese native listeners. The sounds in focus are the retroflex fricative /ʂ/ and alveolo-palatal fricative /ɕ/ in Mandarin. An assimilation experiment and an AXB discrimination experiment were conducted. Finally, this paper will discuss the prediction by Perceptual Assimilation Model [1, 2, 3].

1.1. Background

1.1.1. Mandarin sibilants

There are three sibilants in Mandarin: voiceless alveolar fricative /s/, voiceless so-called retroflex fricative /ʂ/ and voiceless alveolo-palatal fricative /ɕ/ [7, 10, 18]. All of the obstruents in Mandarin are voiceless. The so-called retroflex fricative is also called post-alveolar fricative [12] because it is the upper side, not the underside, of the tip of the tongue that is used to approach the back of the alveolar ridge, which is different from the typical retroflex sounds in Hindi [10]. Phonologically, /s/ and /ʂ/ are contrastive, however, /ɕ/, only appearing when followed by the glides /j/ /ɥ/ or high vowels /i/ /y/, is in complementary distribution with /x/ (the velar fricative) /s/ and /ʂ/ .

Alveolo-palatal consonants in Mandarin have been identified as allophones of velar consonants [5], however, they are also considered as sounds to be an independent series by another study [6]. It is similar to Japanese /ɕ/, which is in complementary distribution with the alveolar /s/ and only appear when followed by glide /j/ and vowel /i/ [16]. All of the sibilants can only appear in the onset position.

There are several studies reporting the various realizations of Mandarin sibilants. /s/ is described as a dental [5, 7] or alveolar [8, 10] fricative, while /ʂ/ is characterized as a retroflex [7] or apical post-alveolar [12] fricative. For /ɕ/, it was characterized as a "palatalized post-alveolar" sound through X-ray and palatograms [11]. A later study [12] using palatograms and linguograms showed that /ɕ/ is a laminal or antero-dorsal alveolar-postalveolar or postalveolar fricative. EPG and EMA data were also collected to examine the place of articulation of Mandarin sibilants [8], reporting that "it is meaningless to characterize /ɕ/ in terms of a distinct place of articulation". The influence of individual difference in various characterizations of Mandarin sibilants has also been reported [14, 15].

1.1.2. Japanese sibilants

There are two voiceless sibilants in Japanese: voiceless alveolar fricative /s/ and voiceless alveolo-palatal fricative /ɕ/ [16, 17]. Compared with English post-alveolar fricative /ʃ/, the constriction of /ɕ/ is longer in distance from front to back although both /s/ and /ɕ/ are laminal fricative. These two sounds are in contrast with each other, however, when appearing before /i/ and /j/, /s/ and /ɕ/ would share the allophone [ɕ], which is called neutralization [9]. All of the sibilants can only appear in the onset position.

1.2. PAM

As a theoretical model, Perceptual Assimilation Model [1, 2] proposes that native speakers of one language perceptually assimilate a pair of sounds from an unknown foreign language into their native phonemic system, according to their phonetic similarities to, or differences from the phonological categories of their native language. During the assimilation process, the learners may categorize the two sounds into different patterns, which predict the speaker's discrimination ability for the sounds.

Two different sounds categorized into two different native phonemes will lead to Two Category assimilation (TC type) in PAM. TC assimilation is the easiest pattern to be recognized. Two sounds categorized into the same phoneme with an equal degree will lead to Single Category assimilation (SC type), which is the most difficult pattern to be recognized. When the two sounds are categorized into the same phoneme with different goodness level, Category Goodness difference (CG type) will be formed.

There are also three other patterns for uncategorized and non-speech sounds. When one or both non-native sounds could not be sufficiently categorized into any given native phoneme, one, or both of them are Uncategorized. If one sound could be categorized as native phoneme when the other could not, it results in Uncategorized- Categorized assimilation (UC type), which is easily recognized. However, if both sounds were uncategorized, Uncategorized-Uncategorized assimilation (UU type) will be formed, whose discrimination depends on the similarity of the sounds to their partially-similar native phones. Finally, when both sounds are treated as non-speech sounds, Non-Assimilable (NA type) will be formed. Recently, PAM was proposed to apply to L2 perceptual learning [3].

1.3. Present study

For Japanese listeners, whose native language has a two-way sibilant contrast, the perception of Mandarin sibilants, with a three-way contrast, is of interest, but not explored yet. While Mandarin alveolo-palatal sibilant /ɕ/ is articulately similar to Japanese /ɕ/, Mandarin retroflex sibilant /ʂ/ is distant from both two Japanese sibilants. The present study primarily aims to investigate the perception of Mandarin sibilants by Japanese naïve listeners. Whether PAM could account for the discrimination performance will also be discussed.

2. EXPERIMENT I

2.1. Stimuli and listeners

One male native speaker (26 years old, from Beijing, China) recorded the stimulus materials. The materials included three nonsense words /asa/ /aɕa/ and /aʂa/. All of these words are in T1 (tone one, the high-level tone). The words were written as Mandarin characters for the speaker and pronounced independently. The stimuli were recorded through a USB microphone (Blue Snowball) and digitized at 44.1Khz with 16 bits accuracy. Besides the stimulus materials, /afa/ and /aha/ were recorded as distractors.

Ten Japanese native speakers (18-23 years old, M=19.8) participated in experiment I. All are university students of Japan. Eight of them were raised in the Greater Tokyo Area (Tokyo, Kanagawa, or Saitama) while the rest was raised in Osaka. None have learned any Chinese language; however, all have learned English from junior-high school.

2.2. Procedures

Experiment I was carried out in a soundproof room through a headphone (MDR-CD900ST). Praat [4] was used as an interface. Before the test, the listeners were asked to familiarized with the task.

The listeners were asked to select the most similar Japanese sound with the stimuli and to rate its category goodness on a 5-point scale. The words /asa/, /asha/, /afa/, and /aha/ were presented as *katakana* for Japanese listeners. Experiment I included 5 stimuli × 2 token × 10 repetitions = 100 trials. The listeners could have a break after every 50 trials and all of the stimuli were presented randomly.

Table 1: The result of Experiment I.

Mandarin Sibilants	Most Common Assimilation	Proportion	Goodness rate (1~5)
/s/	/s/	98%	4.77
/ʂ/	/ɕ/	100%	4.66
/ɕ/	/ɕ/	99%	4.37

2.3. Results

Table 1 shows the result of experiment I. Mandarin /s/ was perceived as Japanese /s/ at the rate of 98% with the mean goodness rate at 4.77. Both Mandarin /ʂ/ and /ɕ/ were perceived as Japanese /ɕ/ in high proportion and mean goodness rates (/ʂ/: 100%, 4.66; /ɕ/: 99%, 4.37). A paired t-test showed no significant

difference between the mean goodness rates of each listener of "/ʂ/ to /ɕ/" and "/ɕ/ to /ɕ/" (t = -1.975, n.s.).

3. EXPERIMENT II

3.1. Stimuli and listeners

One additional female native speaker (26 years old, from Beijing, China) of Mandarin recorded the stimulus materials for Experiment II. The materials are the same nonsense words as Experiment I and recorded through the same equipment.

Another ten Japanese native speakers (22-39 years old, M=26) participated in experiment II. All are university students in Japan. Six were raised in Greater Tokyo Area (Tokyo, Kanagawa, or Saitama) while the rest was raised in other areas. None have learned any Chinese language; however, all have learned English from junior-high school.

3.2. Procedures

The listeners participated in an AXB discrimination test. Each trial included three stimuli, recorded in different tokens, in random order (AAB, ABB, BBA or BAA) and the second stimulus was the same word as the first or the third one but spoken by the different speaker. Then, the listeners were asked to judge which sound (the first or the third one) was more similar to the second one.

Experiment II included 4 pairs (3 sibilant pairs + 1 distractor pair) × 4 AXB orders × 10 repetitions = 160 trails. The sibilant pairs are /s/-/ɕ/, /s/-/ʂ/ and /ʂ/-/ɕ/. The distractor pair is /f/-/h/.

3.3. Results

Fig. 1 presents the result of Experiment II. The mean accuracy rates of Mandarin sibilant contrast were: 96.8% (/s/-/ʂ/, sd=5.5%), 95.3% (/s/-/ɕ/, sd=5.9%), and 68.5% (/ʂ/-/ɕ/, sd=18.6%). The error bars show the standard errors of the mean.

A logistic regression analysis (Gaussian) was conducted for the result of Experiment II. Table 2 shows the summary of the logistic regression, whose dependent variable is accurate rate (correct as 1, incorrect as 0). The independent variable, sibilant pair (/s/-/ʂ/, /s/-/ɕ/, /ʂ/-/ɕ/) is categorical (nominal scale). The /ʂ/-/ɕ/ contrast was significantly different from the results of other two contrasts ($p < .01$). Notably, the discrimination of /ʂ/-/ɕ/ contrast is more difficult for Japanese listeners than other two contrasts. Additionally, the relatively larger standard deviation indicated the greater individual difference for discrimination of /ʂ/-/ɕ/ than other two contrasts.

Figure 1: The result of Experiment II.

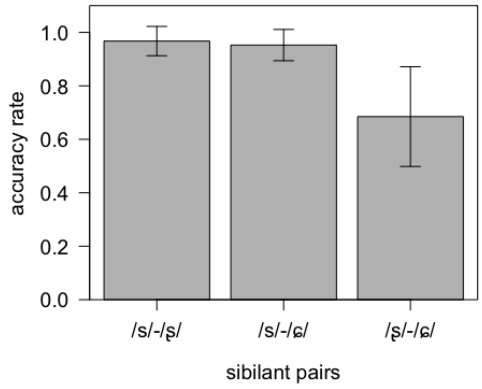

Table 2: The summary of logistic regression analysis of Experiment II.

Independent Variable	Odds Ratio (95% CI)	*P* value
Sibilant pair		
/s/-/ʂ/	reference	-
/s/-/ɕ/	0.99 (0.94-1.03)	0.49
/ʂ/-/ɕ/	0.75 (0.72-0.79)	< 0.001

4. DISCUSSION

4.1. Perception

In the perception of Japanese listeners, Mandarin /s/ was categorized into Japanese /s/ in high goodness rate. Articulately, Mandarin /s/ is similar with Japanese /s/, both of which are alveolar fricatives.

Japanese listeners categorized both Mandarin /ʂ/ and /ɕ/ into Japanese /ɕ/ with similar high goodness rate. Mandarin /ɕ/ is also similar with Japanese /ɕ/ in place and manner of articulation. Although there is no retroflex sound in Japanese, the movement of tongue, approaching the back of the alveolar ridge, when pronouncing Mandarin /ʂ/ is similar with Japanese /ɕ/. Therefore, it is reasonable that Japanese listeners perceived Mandarin /ʂ/ as Japanese /ɕ/.

Discriminations between Mandarin /s/ and other two sibilants are proved to be easy for most Japanese listeners. The contrast of Mandarin /ʂ/-/ɕ/, however, is difficult for them to discriminate.

4.2. Prediction by PAM

Table 3 presents the prediction by PAM according to experiment I and the performance of

listeners in experiment II. According to the prediction by PAM, the assimilation pattern of Mandarin sibilant contrast /s/-/ʂ/ and /s/-/ɕ/ for Japanese native listeners are both TC types. Their discriminations are predicted to be excellent, and listeners indeed performed almost perfectly. As an unfamiliar sound for Japanese listeners, Mandarin retroflex /ʂ/ is considered to be confused with /ɕ/ because of the similarity of articulation. The assimilation pattern of Mandarin sibilant contrast /ʂ/-/ɕ/ is SC type, whose discrimination was predicted to be poor. The result was also consistent with the prediction by PAM.

Table 3: The assimilation pattern and prediction by PAM and the result of the discrimination experiment.

Mandarin Sibilant Contrast	Assimilation Pattern	Prediction	Result
/s/-/ʂ/	TC	Excellent	Excellent
/s/-/ɕ/	TC	Excellent	Excellent
/ʂ/-/ɕ/	SC	Poor	Poor

5. CONCLUSION

The present study investigated the perception of Mandarin three-way sibilant contrast by Japanese listeners without the experience of learning Mandarin. Experiment I revealed that both Mandarin retroflex /ʂ/ and alveolo-palatal /ɕ/ are assimilated into Japanese alveolo-palatal /ɕ/ in the same degree of goodness, leading to a prediction that /ʂ/-/ɕ/ distinction is relatively difficult for the listeners by PAM. Experiment II revealed a lower accuracy rate of distinguishing /ʂ/-/ɕ/ than other two Mandarin sibilant contrasts (/s/-/ʂ/ and /s/-/ɕ/). The results indicated that the discrimination of Mandarin /ʂ/-/ɕ/, whose assimilation pattern is Single Category assimilation, is relatively difficult for Japanese listeners, while other two contrasts can be discriminated with no difficulty. According to the results of experiments, PAM accounts for the pattern of perception of Mandarin sibilants by Japanese listeners with no Mandarin experience.

6. REFERENCES

[1] Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. by. Baltimore, MD: York Press, 171-204.

[2] Best, C. T., McRoberts, G. W., Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109(2), 775-794.

[3] Best, C. T., Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities, In: O.-S. Bohn & M. J. Munro (eds.), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam: John Benjamin, 13-34.

[4] Boersma, P., Weenink, D. (2013). Praat: doing phonetics by computer. http://www.praat.org

[5] Chao, Y. (1968). *A Grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.

[6] Cheng, C. (1973). *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

[7] Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford, UK; New York: Oxford University Press.

[8] Hu, F. (2008). The three sibilants in Standard Chinese. *Proceedings of the 8th Intl. Seminar on Speech Production*.

[9] Labrune, L. (2012). *The phonology of Japanese*. Oxford, UK: Oxford University Press.

[10] Ladefoged, P., Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford, UK; Cambridge, Mass.: Blackwell Publishers.

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

[12] Lee, W. S. (1999). An articulatory and acoustical analysis of the syllable-initial sibilants and approximant in Beijing Mandarin. *Proceedings of the 14th International Congress of Phonetic Sciences*.

[13] Lee, W., Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, 33. 109-112.

[14] Li, C. W. C. (2004). Conflicting notions of language purity: the interplay of archaising, ethnographic, reformist, elitist and xenophobic purism in the perception of Standard Chinese. *Language & Communication*, 24(2), 97-133.

[15] Proctor, M., Lu, L. H., Zhu, Y., Goldstein, L., & Narayanan, S. (2012). Articulation of Mandarin Sibilants: a multi-plane realtime MRI study. *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, Sydney, Australia.

[16] Vance, T. (1987). *An Introduction to Japanese Phonology*. Albany, N.Y.: State University of New York Press.

[17] Vance, T. (2008). *The Sounds of Japanese*. Cambridge, U.K.; New York: Cambridge University Press.

[18] Zhu, X. (2010). *Yu Yin Xue* [Phonetics]. Beijing: The Commercial Press.