DEPARTMENT OF LINGUISTICS AND TRANSLATION

LT6581 PROJECT

Production and Perception of Mandarin Sibilants by Cantonese Speakers

Supervisor: Dr. Lee Wai Sum, Vanti

Postgraduate student: Wong Cheuk Kin

Major: Language Studies

Date: May, 2016

Abstract

The present study investigates the production and perception of the three series of Mandarin sibilants, namely the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], by eight Cantonese-speaking university students. Acoustic analysis and perceptual assessement of the production data were conducted to determine the accuracy and correspondence between the phonetic realization and phonological categorization of the L2 Mandarin sibilants produced by Cantonese speakers. The perceptual data from in a listening test were analyzed for determining the Cantonese speakers' ability in distinguishing the Mandarin sibilants, and they were also compared to the production data for determining the interrelationship between production and perception of L2 Mandarin sounds in Cantonese speakers.

The results show that while the pronunciation of Mandarin sibilants of Cantonese speakers is generally not native-like, most of the Mandarin sibilants are nonetheless identifiable by native speakers of Mandarin. The frication noise patterns of the L2 Mandarin sibilants differ from those of the L1 Cantonese sibilants, suggesting the influence of L1 Cantonese on L2 Mandarin is not significant. The Cantonese speakers' performance is better in perception than in production, indicating a difference in competence between production and perception of the L2 sounds. The present study also discusses its findings in connection with those reported in the previous studies of L2 acquisition of Mandarin sounds and the theories of L2 learning. The findings serve as the foundation for further investigation into the L2 acquisition of Mandarin sounds.

Keywords: acoustic analysis, production and perception, Mandarin sibilants, noise peak and noise range, L2 Mandarin learning

Table of Content

1. Introduction and Background Information

Mandarin, the official language and lingua franca in China, has been taught at the local schools from primary to tertiary level in Hong Kong since 1997. Thus, in addition to English, Mandarin becomes another second language for the majority of Hong Kong students whose first language is Cantonese. Cantonese and Mandarin are two unintelligible dialects of the Chinese language. Due to the differences in the consonants, vowels, and tones in the sound system between Cantonese and Mandarin, pronunciation errors are observable in the Mandarin speech of Cantonese-speaking students, and they have been reported in a number of previous studies (e.g., Ng, 2001; Hon, 2003; Li, 2009; Lee-Wong, 2013; Wu and Su, 2014).

1.1. Sibilant Consonants in Mandarin and Cantonese

In Mandarin, the sibilant consonants are often considered as one type of difficult sounds for Cantonese-speaking students. The sibilant consonants are produced with the high-pitch hissing noise as a distinct acoustic feature (Ladefoged and Johnson, 2015). In both Mandarin and Cantonese, the sibilant consonants include the voiceless affricates and fricatives. According to Lee and Zee (2003), there are three sets of coronal sibilants in Mandarin, namely the denti-alveolar [ʦ, ʦʰ, s], post-alveolar or retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], as presented in Table 1. However, as shown in Table 2, there is only one set of sibilants in Cantonese, namely the alveolar or alveolo-postalveolar [ʦ, ʦʰ, s] (Zee, 1999).

Place / Sibilant types	Dental	Alveolar	Retroflex / Post-alveolar	Alveolo-palatal
Affricate	ʦ ʦʰ		ʈʂ ʈʂʰ	ʨ ʨʰ
Fricative	s		ʂ	ɕ

Table 1. The sibilant consonants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] in Mandarin.

Place / Sibilant types	Alveolar	Post-alveolar
Affricate	ʦ ʦʰ	
Fricative	s	

Table 2. The sibilant consonants [ʦ, ʦʰ, s] in Cantonese.

The denti-alveolar sibilants [ʦ, ʦʰ, s] in Mandarin and the alveolar sibilants [ʦ, ʦʰ, s] in Cantonese are represented with the same IPA symbols, while the sibilants in fact differ in place of articulation between the two languages. For the Mandarin [ʦ, ʦʰ, s], the contact is made between the tongue tip and/or blade and the alveolar ridge extending forward to the dental area (Lee and Zee, 2003). As for the Cantonese [ʦ, ʦʰ, s], the tongue blade is mainly used to make contact with the alveolar ridge, with the contact extending backward to the post-alveolar area (Zee, 1999). Thus, the sibilants [ʦ, ʦʰ, s] are produced in a more forward place of the vocal tract in Mandarin than Cantonese.

The other two sets of sibilants, the post-alveolar [ʈʂ, ʈʂʰ, ʂ] and alveolo-palatal [ʨ, ʨʰ, ɕ], in Mandarin are not found in Cantonese. The Mandarin post-alveolar sibilants [ʈʂ, ʈʂʰ, ʂ], conventionally referred to as retroflexes, are made with no tip curling, but with the tongue tip retracting towards to the post-alveolar area in the present-day young generation (Lee and Zee, 2003). As for the Mandarin alveolo-palatal sibilants [ʨ, ʨʰ, ɕ], both the tongue blade and front dorsum are involved in the articulation, making extensive contact on the palate extending from the pre-palatal area to the post-alveolar area (Lee and Zee, 2003). Thus, in terms of articulation, the Cantonese alveolar [ʦ, ʦʰ, s] are similar to the alveolo-palatal [ʨ, ʨʰ, ɕ] in Mandarin.

Furthermore, according to Lee and Zee (2003), the three sets of sibilants in Mandarin are adjacent to different allophones of the vowel phoneme /i/ when they

occur in open CV syllables. While the alveolo-palatal [tɕ, tɕʰ, ɕ] are followed by the dorsal high front vowel [i], the denti-alveolar [ʦ, ʦʰ, s] and retroflex [ʈʂ, ʈʂʰ, ʂ] are followed by the respective homorganic apical vowels [ɿ] and [ʅ], the two allophonic variants of the vowel /i/. Different from Mandarin, Cantonese has no variants of apical vowel for the vowel /i/, and the only set of alveolar sibilants [ʦ, ʦʰ, s] in Cantonese are followed by the high front vowel [i] (Zee, 1999). Thus, the Cantonese [ʦ, ʦʰ, s] are also similar to the Mandarin [tɕ, tɕʰ, ɕ] in terms of the adjacent vowel.

Acoustically, the sibilant consonants are distinct in the pattern of noise energy distribution, which is specifically concerned with the frequency values of the noise peak and the noise range - the two primary acoustic attributes to the place of articulation of the sibilants (Heiz and Stevens, 1961; Behrens and Blumstein, 1988; Evers, Reetz, and Lahiri, 1998; Nowak, 2006). According to some recent acoustic studies of the three sets of sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [tɕ, tɕʰ, ɕ] in Mandarin (Chung, 2009; Chung and Si, 2009; Lee, 2011; Lee, Zhang, and Li, 2014; Li and Gu, 2015; Wong, 2015), the frequency value of the noise peak is the highest for denti-alveolar [ʦ, ʦʰ, s] and the lowest for the retroflex [ʈʂ, ʈʂʰ, ʂ], with the alveolo-palatal [tɕ, tɕʰ, ɕ] in between. In general, the frequency value for the noise peak of sibilants decreases when the sibilants are articulated with the tongue moving backward in the mouth (Chung, 2009; Lee et al., 2014). In Mandarin, while the constriction locations for the alveolo-palatal [tɕ, tɕʰ, ɕ] and retroflex [ʈʂ, ʈʂʰ, ʂ] sibilants are both near to the post-alveolar area, the frequency value for the noise peak is much lower for [ʈʂ, ʈʂʰ, ʂ] than [tɕ, tɕʰ, ɕ] (Svantesson, 1986; Chung, 2009; Chung and Si, 2009; Lee, 2011; Lee, Zhang, and Li, 2014; Li and Gu, 2015; Wong, 2015). It is considered that the lowering of the frequency value for the noise peak of [ʈʂ, ʈʂʰ, ʂ] may be caused by the occurrence of a large sublingual cavity (Lee, 2011), and/or the lengthening of the front cavity (Chung, 2009; Lee et al., 2014; Li and Gu, 2015), due

to the backward movement of the constriction location during the retroflex articulation of [tʂ, tʂʰ, ʂ].

1.2. Studies of L2 Acquisition of Mandarin Sounds

In view of the non-occurrence of the alveolo-palatal [tɕ, tɕʰ, ɕ] and retroflex [tʂ, tʂʰ, ʂ] sibilants in Cantonese and the difference in place of articulation between the Mandarin denti-alveolar [ʦ, ʦʰ, s] and Cantonese alveolar [ʦ, ʦʰ, s], it is expected that Cantonese speakers have difficulty in producing and distinguishing the three place categories of sibilants in Mandarin. Wong (2015) is an acoustic study of the production of the three sets of Mandarin sibilants, [ʦ, ʦʰ, s], [tʂ, tʂʰ, ʂ], and [tɕ, tɕʰ, ɕ], by Cantonese-speaking Mandarin learners. The frequency values for the noise peak and noise range of the test sibilants from two male and two female university students who had taken an elementary Mandarin course were measured and compared with those from a Mandarin speaker. Table 3 presents the correct and incorrect percentages of the Mandarin sibilants produced by the four Cantonese speakers in Wong (2015).

Classification / **Sibilant types**	**'Correct'**	**'Incorrect'**	
		➔ Other sibilants	**➔ New form**
Denti-alveolar [ʦ, ʦʰ, s]	8/72 (11.11%)	27/72 (37.50%)	37/72 (51.39%)
Retroflex [tʂ, tʂʰ, ʂ]	50/72 (69.44%)	11/72 (15.28%)	11/72 (15.28%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	18/72 (25%)	9/72 (12.5%)	45/72 (62.5%)
Overall	**35.19%**	**21.75%**	**43.06%**

Table 3. Correct and incorrect percentages of the three sets of Mandarin sibilants produced by four Cantonese speakers (data from Wong, 2015; a total of 216 tokens, with 72 tokens for each place category of the sibilants).

As presented in Table 3, only 35% of a total of 216 test Mandarin sibilants from the four Cantonese speakers are correctly produced, while the remaining 65% of the test sibilants are either mispronounced as other sibilants (21.75%) or classified as a 'new form' (43.06%) which is not the same as any one of the three place categories of sibilants in Mandarin. For the correct cases, 69.44% is for the retroflex [tʂ, tʂʰ, ʂ], which is much higher than the correct rates of 25% for the alveolo-palatal [tɕ, tɕʰ, ɕ] and 11.11% for the denti-alveolar [ts, tsʰ, s]. Correspondingly, for the incorrect cases, the error rate is noticeably lower for [tʂ, tʂʰ, ʂ] (30.56%) than for [ts, tsʰ, s] (88.89%) and [tɕ, tɕʰ, ɕ] (75%). Thus, Wong concludes that the production of the Mandarin denti-alveolar [ts, tsʰ, s] and alveolo-palatal [tɕ, tɕʰ, ɕ] are more difficult than the retroflex [tʂ, tʂʰ, ʂ] for Cantonese speakers.

Wong (2015) made a comparison of the production data for Cantonese speakers with those for the other three groups of Mandarin beginners with other L1, including Korean, Japanese, and Vietnamese, reported in Chung and Si (2009). The results of comparison show that the four groups of L2 speakers share some error patterns for the production of the Mandarin sibilants, such as (i) the mispronunciation of the denti-alveolar [ts, tsʰ, s] as the retroflex [tʂ, tʂʰ, ʂ] or alveolo-palatal [tɕ, tɕʰ, ɕ], and (ii) the mispronunciation of [tʂ, tʂʰ, ʂ] as [tɕ, tɕʰ, ɕ] or in vice versa. However, there are some other error patterns, such as the mispronunciation of [tʂ, tʂʰ, ʂ] and [tɕ, tɕʰ, ɕ] as [ts, tsʰ, s], which are observed in Cantonese speakers but not in the other three groups of L2 speakers. The differences across the four groups of L2 speakers may be taken as an indication of the language-specific effect of L1 on L2 acquisition.

1.3. Theories of L2 Acquisition

The L2 production data in Wong (2015) support the speech learning model proposed by Flege (1995) regarding the acquisition of L2 sounds. The model predicts

that the ‘identical phones’, which are the same sounds in both L1 and L2, and the ‘new phones’, which only exist in L2 but not L1, are easier to be acquired than the ‘similar phones’, which are the similar sounds between L1 and L2. In Wong (2015), the Mandarin denti-alveolar [ʦ, ʦʰ, s], which are similar to the Cantonese alveolar [ʦ, ʦʰ, s] and considered as ‘similar phones’, are the most difficult for Cantonese speakers. The Mandarin retroflex [ʈʂ, ʈʂʰ, ʂ] and alveolo-palatal [ʨ, ʨʰ, ɕ], which are non-occurring in Cantonese and considered as ‘new phones’, are relatively easy to be produced by Cantonese speakers. However, between the two sets of ‘new phones’, the retroflex [ʈʂ, ʈʂʰ, ʂ] are much easier to be produced than the alveolo-palatal [ʨ, ʨʰ, ɕ] for Cantonese speakers. In view of the fact that the Mandarin [ʨ, ʨʰ, ɕ] are produced with a large constriction extending from the pre-palatal area forwards to the alveolar ridge and the constriction area during the alveolar [ʦ, ʦʰ, s] in Cantonese may extend to the post-alveolar area, Wong (2015) proposed that the category of ‘new phones’ can be further divided into two types, namely ‘new similar phones’ and ‘new non-similar phones’, and Wong’s data suggest that the ‘new non-similar phones’ [ʈʂ, ʈʂʰ, ʂ] are easier to be produced than the ‘new similar phones’ [ʨ, ʨʰ, ɕ].

There are other studies of the production of Mandarin sibilants in L2 addressing the theory regarding the effect of L1 transfer on L2 acquisition (Ng, 2001; Hon, 2003; Chung, 2009; Chung and Si, 2009; Lee-Wong, 2013). On the basis of the similarity in articulation between the Cantonese [ʦ, ʦʰ, s] and the Mandarin [ʨ, ʨʰ, ɕ], Ng (2001) argues that it is a case of negative transfer, where Cantonese speakers produce the Mandarin [ʨ, ʨʰ, ɕ] as Cantonese [ʦ, ʦʰ, s]. The observations of the Mandarin speech of Cantonese-speaking learners in Lee-Wong (2013) are basically in agreement with Ng’s (2001) view, but in addition to the confusion between [ʦ, ʦʰ, s] and [ʨ, ʨʰ, ɕ], such as [ʦʅ] 自 ‘from’ → [ʨi], the confusion between [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ], such as [ʈʂʰʅ] 遲 ‘late’ → [ʨʰi] is observed. Hon (2003) also reports that a

negative L1 transfer for Cantonese speakers to use the single set of sibilants [ʦ, ʦʰ, s] in Cantonese to replace all the three sets of sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] in Mandarin.

The effect of L1 transfer on L2 is also reported for Mandarin learners with other L1, such as English, Korean, Japanese, Vietnamese, and Thai (Chung and Si, 2009; Chung, 2009). Generally, it is common for the subjects to use the L1 sounds to replace the Mandarin sibilants. According to Chung and Si (2009), and Chung (2009), similar case of L1 transfer is observed in the Korean and Japanese speakers, where the Korean [s] or [s*] and Japanese [s] are used to replace the Mandarin [s], and the Korean [ʃ] and Japanese [ʃ] to replace the Mandarin [ɕ]. For Thai speakers, they use the Thai [s] to replace the Mandarin [ɕ] which is non-occurring in Thai. For English speakers, they use the English syllable-final cluster [ts] as in [kæts] 'cats' to replace the Mandarin [ʦ, ʦʰ]. The effect of negative transfer is noticeably apparent for Vietnamese speakers, where they use the Vietnamese [s] to substitute for all the Mandarin fricatives [s, ʂ, ɕ].

1.4. Perception Studies of L2 Acquisition of Mandarin Sounds

In addition to the production of Mandarin sibilants in L2, there are some studies of the perception of Mandarin sibilants in L2 speakers. Lai (2009) carried out a perceptual test of the Mandarin affricates, including the denti-alveolar [ʦ, ʦʰ], retroflex [ʈʂ, ʈʂʰ], and alveolo-palatal [ʨ, ʨʰ], for 10 Malay- and 10 Burmese-speaking learners of Mandarin Chinese. The results show that both groups of subjects did not perform very well in discrimination between the three sets of affricates. Comparatively, their performance is slightly better in discriminating between [ʦ, ʦʰ] and [ʨ, ʨʰ], and between [ʈʂ, ʈʂʰ] and [ʨ, ʨʰ] than between [ʦ, ʦʰ] and [ʈʂ, ʈʂʰ]. And,

the discrimination is better for the unaspirated affricates than the aspirated equivalents.

The findings in Lai (2009) in general do not support the Speech Learning Model proposed by Flege (1995). In Malay, there is only one pair of the voiced and voiceless post-alveolar affricates [tʃ, dʒ], while in Burmese, there are three post-alveolar affricates [tʃʰ, tʃ, dʒ] with the contrast in aspiration and voicing. The affricates in Malaya and Burmese are not similar to any one of the three place categories of affricates [ʦ, ʦʰ], [ʧ, ʧʰ], and [ʨ, ʨʰ] in Mandarin. According to Flege's Speech Learning Model, the Mandarin affricates are all classified as 'new phones' for both the Malay and Burmese speakers, and thus presumably all the Mandarin affricates are easy to be identified by the two groups of L2 speakers. However, the performance of the subjects in discrimination of the Mandarin affricates is poor and the identification rate varies for the affricates in different place categories and aspiration.

1.5. Purpose of Study

The present research project is a further study of Wong (2015) on the acquisition of Mandarin sibilants by L2 Cantonese-speaking learners. Wong (2015) is a preliminary study of the production of Mandarin sibilants, based on a small amount of speech data collected from four Cantonese speakers. Furthermore, due to the unavailability of the production data on the sibilants in L1 Cantonese for the speakers of the study, Wong has not looked deeply into the negative transfer from L1 to L2. Also, only production data were obtained in Wong (2015). The perceptual ability of Cantonese speakers to distinguish the three place categories of Mandarin sibilants as well as the relationship between production and perception of Mandarin sibilants for Cantonese speakers are still unknown. All of these motivate the present research

project to look further into the L2 acquisition of Mandarin sibilants by Cantonese speakers in both the production and perception aspects.

The present research project investigates the production and perception of the three place categories of Mandarin sibilants, namely the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], by Cantonese-speaking university students in Hong Kong. For the production part, the Mandarin sibilants produced by Cantonese speakers are recorded and analyzed by performing acoustic analysis and also perceptually assessed by a group of native speakers of Mandarin for determining the correspondence between the phonetic realization and phonological categorization of the sibilants. The speech data on the sibilant consonants [ʦ, ʦʰ, s] in Cantonese are also collected from Cantonese speakers for determining the effect of L1 transfer, specifically the substitution of L1 Cantonese sibilants for the L2 Mandarin sounds. For the perception part, the identification scores for Cantonese speakers in a listening test of the Mandarin sibilants produced by a native Mandarin speaker are obtained for determining their perceptual ability of the distinction of the three place categories of sibilants in Mandarin. The error patterns in perception are further compared with those in production for determining the relationship between production and perception in L2 Mandarin sibilant acquisition for Cantonese speakers.

2. Methodology

In the present research project, both production experiment and perception experiment were carried out to investigate the production and perception of the three place categories of Mandarin sibilants, i.e., the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], for Cantonese speakers.

2.1. Production Experiment

2.1.1. Subjects

In the production experiment, a group of eight Cantonese-speaking university students, four male and four female, aged 18 to 23, were invited to provide speech samples. All the speakers have taken an elementary Mandarin course at the City University of Hong Kong and have the knowledge of the sound system and the pronunciation of the sounds of Mandarin. They took part in an individual audio recording session to produce two sets of monosyllabic words that contain the sibilant consonants in Mandarin and Cantonese.

2.1.2. Test materials

Table 4 presents the two sets of monosyllabic words, Set I and Set II, used for the production experiment of the present study. As can be seen, Set I consists of nine Mandarin CV monosyllabic words, with the sibilants, denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], in the word-initial position followed by one of the three vowels [i, ɿ, ʅ] which are considered as the allophonic variants of the vowel phoneme /i/ in Mandarin. Set II consists of the test monosyllabic words in Cantonese, which are also in CV syllable structure, with the alveolar sibilants [ʦ, ʦʰ, s] in the word-initial position followed by the vowel [i]. The two sets of test words were used for determining (i) the similarities and differences between the sibilant

consonants in the two languages produced by Cantonese speakers and (ii) the effect of transfer from L1 Cantonese on L2 Mandarin.

The selected test words in both Mandarin and Cantonese are meaningful, and they are familiar to all Cantonese speakers in this project. The two sets of test words were presented in Chinese characters without any pinyin letters on a separate list. Three repetitions of each test word were pseudo-randomized on the lists, in order to avoid the same words appearing in successive sequential order. A total of 288 test tokens (12 test words x 3 repetitions x 8 subjects) were recorded of the Cantonese speakers for subsequent analysis.

Set I: Mandarin			**Set II: Cantonese**
[tsɿ˥] 資 'capital'	[tʂɻ˥] 知 'know'	[tɕi˥] 基 'base'	[tsɿ˥] 資 'capital'
[tsʰɿ˥] 疵 'flaw'	[tʂʰɻ˥] 吃 'eat'	[tɕʰi˥] 七 'seven'	[tsʰɿ˥] 疵 'flaw'
[sɿ˥] 司 'in charge of'	[ʂɻ˥] 失 'lose'	[ɕi˥] 希 'hope'	[sɿ˥] 司 'in charge of'

Table 4. Two sets of test monosyllabic words, Set I in Mandarin and Set II in Cantonese, used for investigation.

2.1.3. Data collection and analysis

The audio recordings were carried out in the sound-treated laboratory of the Department of Linguistics and Translation at the City University of Hong Kong. A high quality recorder of Marantz (Model PM661) was used for the recordings with the sampling rate of 44 kHz.

The recorded speech data from Cantonese speakers were analyzed through performing acoustic analysis and perceptual assessment. Using the speech analysis software, Praat (version 6017, 32 bits edition), FFT and LPC spectral analysis was

carried out for the measurements of the frequency values for the two major acoustic attributes, the noise peak and the range of noise distribution, to the test sibilants. The measured frequency data for Cantonese speakers were compared with those for a native female Mandarin speaker who was a Mandarin teacher and an examiner of the National Putonghua Proficiency Test participated in Wong (2015) for determining the error patterns of the production of Mandarin sibilants in Cantonese speakers.

The Mandarin sibilants produced by Cantonese speakers were also assessed perceptually by a group of ten native Mandarin speakers, five male and five female, who were studying Chinese Language or Linguistics at the universities in Hong Kong and had the phonetic and linguistic knowledge of Mandarin Chinese. The Mandarin listeners were asked on the basis of their perception to evaluate and identify the Mandarin sibilants in the test words produced by eight Cantonese speakers. Three tokens of each of the nine test words in Set I (Table 4) from the same Cantonese speaker were randomized in a block. A total of 27 words (9 test words x 3 tokens) in each of the eight blocks for eight Cantonese speakers were played one time to the Mandarin listeners on a computer at a comfortable volume level through earphones. After hearing each word token, four choices presented on an answer sheet were given to the listeners for selection. The choices consisted of (i) three words with the syllable-initial sibilants contrasting in the three place categories, i.e., denti-alveolar, retroflex, and alveolo-palatal (e.g., [tsɿ˩], [tʂɻ˩], and [tɕi˩]), and (ii) 'NA', i.e., none of the three. A total of 2160 responses (27 words x 8 blocks x 10 listeners) were obtained from ten Mandarin listeners. The responses were analyzed for a 'perceptual' set of correct and incorrect rates of the production of Mandarin sibilants for Cantonese speakers. The rates were used for comparing with the 'production' set of correct and incorrect rates based on the frequency data of the sibilants for determining

the correspondence between the phonetic realization and phonological categorization of the Mandarin sibilants produced by Cantonese speakers.

2.2. Perception Experiment

The eight Cantonese speakers took part in a perception experiment after doing the recordings. The task for the speakers was to identify the test words that contain the different sibilant consonants in Mandarin produced by a native female Mandarin teacher in Wong (2015). The stimuli were the nine test words in Set I as presented in Table 4, which were from the natural speech of the Mandarin speaker.

There were a total of 45 stimuli, which consisted of three copies of each of the nine test words from the Mandarin speaker, used for the perception test. The stimuli were divided into three blocks, with the words [tsɿ, tʂɻ, tɕi] containing the different initial unaspirated affricates in Block I, and the words [tsʰɿ, tʂʰɻ, tɕʰi] with the different initial aspirated affricates in Block II, and the words [sɿ, ʂɻ, ɕi] with the different initial fricatives in Block III. Thus, the three words in each block contrast in the place categories, i.e., denti-alveolar, retroflex, and alveolo-palatal, of the initial sibilants.

The stimuli of the three blocks were played one by one to Cantonese subjects on a computer at a comfortable volume level through earphones. Each stimulus was played one time to the subjects. The subjects were asked to identify each stimulus that they just heard each time by choosing one of the three words tested in a particular block shown on an answer sheet. There was no time limit for the subjects to choose the answers, though the subjects were instructed to give response immediately after hearing each stimulus. A total 360 responses (45 stimuli x 8 subjects) were obtained from eight Cantonese subjects. The responses were analyzed for the correct and incorrect rates for the three initial sibilants in each block for determining the

perceptual ability of Cantonese speakers to distinguish the three place categories of sibilants in Mandarin.

3. Results

The results of the production and perception experiments of this study are presented in different parts of this section. The first part presents the noise frequency data, in terms of the noise peak and noise energy range, for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] produced by the eight Cantonese speakers, in comparison of (i) those for a native Mandarin speaker for determining the error patterns in Cantonese speakers, (ii) the frequency data on the Cantonese sibilants [ʦ, ʦʰ, s] produced by the Cantonese speakers for determining the effect of transfer from L1 Cantonese to L2 Mandarin in sibilant production, and (iii) the results of the perceptual assessment performed by native Mandarin speakers for determing the correspondence between the phonetic realization and phonological categorization of the Mandarin sibilants from Cantonese speakers. The second part presents the results of perception of the Mandarin sibilants by Cantonese speakers for determining the perceptual ability to distinguish the different Mandarin sibilants for Cantonese speakers.

3.1. Results of Production Experiment

3.1.1. Noise frequency data of the Mandarin sibilants for Cantonese speakers

Tables 6-13 present the frequency data on the noise distribution, in terms of the noise peak and noise range, for the Mandarin sibilants, the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], produced by each of the eight Cantonese speakers, four male (Cantonese Male 1-4) and four female (Cantonese Female 1-4). For each sibilant, the frequency value of the noise peak or noise range is the mean of three tokens from a particular speaker. For comparison purposes, Table 5 presents the mean frequency data also based on three tokens of each of the Mandarin sibilants for a native Mandarin female speaker in Wong (2015). The speech samples

of the Mandarin speaker in Wong (2015) were re-analyzed in the present study, as the sampling rate of 22 kHz was used in Wong (2015), but 44 kHz in the present study.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	7,960	8,107	12,981
	[ʦʰ]	10,098	7,885	12,046
	[s]	9,772	8,082	11,283
	Mean	**9,145**	**8,025**	**12,103**
Retroflex	[ʈʂ]	3,608	2,001	10,126
	[ʈʂʰ]	3,626	2,322	9,732
	[ʂ]	3,711	1,854	10,126
	Mean	**3,648**	**2,059**	**9,994**
Alveolo-palatal	[ʨ]	6,685	6,081	10,766
	[ʨʰ]	7,250	5,694	9,683
	[ɕ]	6,966	6,113	10,027
	Mean	**6,967**	**5,963**	**10,158**

Table 5. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for a native Mandarin speaker.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	9,269	7,780	11,310
	[ʦʰ]	8,183	6,322	10,634
	[s]	9,557	7,599	10,664
	Mean	**9,003**	**7,234**	**10,869**
Retroflex	[ʈʂ]	6,423	3,062	11,100
	[ʈʂʰ]	8,324	5,977	10,198
	[ʂ]	7,201	2,311	11,160
	Mean	**7,316**	**3,783**	**10,819**
Alveolo-palatal	[ʨ]	8,844	7,164	11,310
	[ʨʰ]	9,346	6,953	10,904
	[ɕ]	8,896	7,134	10,514
	Mean	**9,029**	**7,084**	**10,909**

Table 6. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Male 1.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	8,557	7,314	8,118
	[ʦʰ]	8,560	7,564	11,205
	[s]	8,970	7,481	11,997
	Mean	**8,696**	**7,453**	**10,440**
Retroflex	[ʈʂ]	8,935	7,717	11,872
	[ʈʂʰ]	8,479	7,203	11,150
	[ʂ]	8,535	7,495	12,108
	Mean	**8,650**	**7,472**	**11,710**
Alveolo-palatal	[ʨ]	8,586	7,287	10,899
	[ʨʰ]	8,184	7,120	12,275
	[ɕ]	8,252	7,106	11,761
	Mean	**8,341**	**7,171**	**11,645**

Table 7. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Male 2.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	9,231	7,036	12,887
	[ʦʰ]	8,587	7,398	13,123
	[s]	7,964	2,478	13,845
	Mean	**8,594**	**5,637**	**13,285**
Retroflex	[ʈʂ]	7,359	3,673	13,706
	[ʈʂʰ]	9,442	5,524	13,067
	[ʂ]	8,782	7,312	14,351
	Mean	**8,528**	**5,503**	**13,708**
Alveolo-palatal	[ʨ]	9,529	7,484	14,743
	[ʨʰ]	8,883	7,253	13,716
	[ɕ]	8,428	7,189	12,063
	Mean	**8,947**	**7,309**	**13,507**

Table 8. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Male 3.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	9,562	7,379	11,703
	[ʦʰ]	10,328	7,422	12,169
	[s]	9,507	7,697	11,300
	Mean	**9,799**	**7,499**	**11,724**
Retroflex	[ʈʂ]	5,688	2,378	10,961
	[ʈʂʰ]	6,290	2,230	10,474
	[ʂ]	3,968	2,251	10,664
	Mean	**5,315**	**2,286**	**10,700**
Alveolo-palatal	[ʨ]	9,929	7,538	11,764
	[ʨʰ]	8,168	6,829	11,745
	[ɕ]	10,033	7,062	11,512
	Mean	**9,377**	**7,143**	**11,674**

Table 9. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Male 4.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	9,194	6,493	10,375
	[ʦʰ]	8,360	5,761	10,629
	[s]	5,550	4,103	10,569
	Mean	**7,701**	**5,452**	**10,524**
Retroflex	[ʈʂ]	8,567	6,284	10,226
	[ʈʂʰ]	5,624	4,238	10,032
	[ʂ]	5,291	3,506	9,181
	Mean	**6,494**	**4,676**	**9,813**
Alveolo-palatal	[ʨ]	8,335	6,433	9,778
	[ʨʰ]	9,191	6,418	11,032
	[ɕ]	5,929	5,358	10,121
	Mean	**7,818**	**6,070**	**10,310**

Table 10. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Female 1.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	9,511	6,466	11,636
	[ʦʰ]	6,505	5,216	9,968
	[s]	9,557	6,652	11,181
	Mean	**8,524**	**6,111**	**10,928**
Retroflex	[ʈʂ]	9,429	7,135	11,107
	[ʈʂʰ]	8,184	6,717	11,017
	[ʂ]	9,168	6,851	10,808
	Mean	**8,927**	**6,901**	**10,977**
Alveolo-palatal	[ʨ]	9,334	6,254	10,196
	[ʨʰ]	9,097	6,702	11,330
	[ɕ]	9,477	6,418	10,778
	Mean	**9,303**	**6,458**	**10,768**

Table 11. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Female 2.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	8,541	7,867	11,853
	[ʦʰ]	9,413	7,254	11,301
	[s]	9,585	7,583	11,928
	Mean	**9,180**	**7,568**	**11,694**
Retroflex	[ʈʂ]	9,063	7,463	11,301
	[ʈʂʰ]	8,979	7,344	10,629
	[ʂ]	6,484	6,179	11,346
	Mean	**8,175**	**6,995**	**11,092**
Alveolo-palatal	[ʨ]	8,613	6,687	11,555
	[ʨʰ]	9,501	7,015	11,808
	[ɕ]	8,994	6,762	11,898
	Mean	**9,036**	**6,821**	**11,754**

Table 12. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Female 3.

Target sibilants		Noise peak	Noise range	
			Minimum value	Maximum value
Denti-alveolar	[ʦ]	9,258	7,538	11,584
	[ʦʰ]	9,502	7,029	10,922
	[s]	9,390	8,906	12,392
	Mean	**9,383**	**7,824**	**11,633**
Retroflex	[ʈʂ]	5,245	2,837	9,417
	[ʈʂʰ]	4,532	2,972	9,883
	[ʂ]	3,669	2,537	9,522
	Mean	**4,482**	**2,782**	**9,607**
Alveolo-palatal	[ʨ]	6,886	4,955	10,439
	[ʨʰ]	6,144	5,481	9,357
	[ɕ]	7,233	5,812	10,198
	Mean	**6,754**	**5,416**	**9,998**

Table 13. Mean frequency values (in Hz) of the noise peak and noise range for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] for Cantonese Female 4.

As shown in Table 5 for the native Mandarin speaker, it can be seen that the noise patterns for the three place categories of Mandarin sibilants differ from each other. In terms of the noise peak, the frequency is the highest for the denti-alveolar [ʦ, ʦʰ, s] (7,960 Hz to 10,098 Hz), followed by the alveolo-palatal [ʨ, ʨʰ, ɕ] (6,685 Hz to 7,250 Hz) and then the retroflex [ʈʂ, ʈʂʰ, ʂ] (3,608 Hz to 3,711 Hz) in descending order. Concerning the noise range, the maximum value is also the highest for the denti-alveolar [ʦ, ʦʰ, s] (11,283 Hz to 12,981 Hz), but the difference between the alveolo-palatal [ʨ, ʨʰ, ɕ] (9,732 Hz to 10,126 Hz) and retroflex [ʈʂ, ʈʂʰ, ʂ] (9,683 Hz to 10,766 Hz) is minimal. The minimum value of the noise range however is clearly different among the three types of sibilants, where the descending order of the frequency value is also [ʦ, ʦʰ, s] (7,885 Hz to 8,107 Hz) > [ʨ, ʨʰ, ɕ] (5,694 Hz to 6,113 Hz) > [ʈʂ, ʈʂʰ, ʂ] (1,854 Hz to 2,322 Hz). Similar patterns of the acoustic differences among the three types of Mandarin sibilants were also reported in a number of previous studies (Chung, 2009; Chung and Si, 2009; Lee, 2011; Lee,

Zhang, and Li, 2014; Li and Gu, 2015). Therefore, the noise peak and the minimum value of noise range (minimum value, henceforth) are considered as the two acoustic attributes to the place of articulation of the three types of sibilants in Mandarin, and they are taken for the subsequent comparison of the Mandarin sibilants produced by Cantonese speakers with those from the Mandarin speaker.

A comparison of the frequency data on the three types of Mandarin sibilants for Cantonese Male 1 (Table 6) with those from the Mandarin speaker (Table 5) shows that this Cantonese speaker produces the denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ] sibilants with the noise peak (9,003 Hz and 9,029 Hz) and the minimum value (7,234 Hz and 7,084 Hz) similar to those for the Mandarin speaker's denti-alveolar [ʦ, ʦʰ, s] (noise peak: 9,145 Hz; minimum value: 8,025 Hz). The data indicate that Cantonese Male 1 merges the denti-alveolar [ʦ, ʦʰ, s], and alveolo-palatal [ʨ, ʨʰ, ɕ] sibilants into [ʦ, ʦʰ, s]. As for the retroflex [ʈʂ, ʈʂʰ, ʂ], the noise peak (7,316 Hz) for Cantonese Male 1 is similar to the noise peak for Mandarin speaker's alveolo-palatal [ʨ, ʨʰ, ɕ] (6,967 Hz), while the minimum value (3,783 Hz) for Cantonese Male 1 is similar to the Mandarin speaker's [ʈʂ, ʈʂʰ, ʂ] (2,059 Hz). Thus, in terms of both the noise peak and minimum value, it is considered that Cantonese Male 1's [ʈʂ, ʈʂʰ, ʂ] become a 'new form' with the noise pattern similar to both the Mandarin [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ] which cannot be classified to any one of the three types of sibilants in Mandarin.

The error patterns of the Mandarin sibilants production for Male Cantonese 1 are also observed in Cantonese Male 4 (Table 9). That is, (i) the denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ] are produced with the noise peak (9,799 Hz and 9,377 Hz) and minimum value (7,499 Hz and 7,143 Hz) similar to those for the Mandarin speaker's [ʦ, ʦʰ, s] (9,145 Hz and 8,025 Hz); and (ii) the retroflex [ʈʂ, ʈʂʰ, ʂ]'s minimum value (2,286 Hz), but not the noise peak (5,315 Hz), is similar to the

Mandarin speaker's [tʂ, tʂʰ, ʂ] (noise peak: 3,648 Hz; minimum value: 2,059 Hz). Thus, for Cantonese Male 4, it is also considered that the sibilants [ts, tsʰ, s] and [tɕ, tɕʰ, ɕ] merge into [ts, tsʰ, s], and the sibilants [tʂ, tʂʰ, ʂ] become a 'new form' which cannot be classified to any type of sibilants in Mandarin.

For Cantonese Male 2-3 and Cantonese Female 1-3, the three types of Mandarin sibilants are produced with similar noise frequency pattern, indicating that the different place categories of sibilants are not distinguishable in their speech. As shown in Table 7 for Cantonese Male 2, the frequency values of noise peak and minimum value for the sibilants [ts, tsʰ, s] (8,696 Hz and 7,453 Hz), [tʂ, tʂʰ, ʂ] (8,650 Hz and 7,472 Hz), and [tɕ, tɕʰ, ɕ] (8,341 Hz and 7,171 Hz) are similar, and they are close to those for the denti-alveolar [ts, tsʰ, s] (9,145 Hz and 8,025 Hz) for the Mandarin speaker (Table 5). The data indicate that the three types of Mandarin sibilants merge into one as [ts, tsʰ, s] for Cantonese Male 2. For Cantonese Male 3 (Table 8), Cantonese Female 2 (Table 11), and Cantonese Female 3 (Table 12), the frequency values of the noise peak for [ts, tsʰ, s], [tʂ, tʂʰ, ʂ], and [tɕ, tɕʰ, ɕ] are in the range of 8,175 Hz to 9,303 Hz, which are similar to the frequency value of the noise peak for the Mandarin speaker's [ts, tsʰ, s] (9,145 Hz). But, in terms of the minimum value, these Cantonese speakers' [ts, tsʰ, s], [tʂ, tʂʰ, ʂ], and [tɕ, tɕʰ, ɕ] have the frequency value in the range of 5,503 Hz to 7,309 Hz similar to the Mandarin speaker's [ts, tsʰ, s] (8,025 Hz) or [tɕ, tɕʰ, ɕ] (5,963 Hz). The frequency data indicate that the three types of Mandarin sibilants for these Cantonese speakers are produced as the denti-alveolar [ts, tsʰ, s] or a 'new form' associated with the noise pattern similar to those for the denti-alveolar [ts, tsʰ, s] and alveolo-palatal [tɕ, tɕʰ, ɕ].

For Cantonese Female 1 (Table 10), the frequency values of the minimum value for [ts, tsʰ, s] (5,452 Hz), [tʂ, tʂʰ, ʂ] (4,676 Hz), and [tɕ, tɕʰ, ɕ] (6,070 Hz) are similar to that for the Mandarin speaker' [tɕ, tɕʰ, ɕ] (5,963 Hz). But, in terms of the

frequency value of the noise peak, Cantonese Female 1's [ʦ, ʦʰ, s] (7,701 Hz), [ʈʂ, ʈʂʰ, ʂ] (6,494 Hz) and [ʨ, ʨʰ, ɕ] (7,818 Hz) are similar to the Mandarin speaker's [ʨ, ʨʰ, ɕ] (6,967 Hz) or [ʦ, ʦʰ, s] (9,145 Hz). Thus, for this speaker, it may be considered that the three types of Mandarin sibilants are produced as the alveolo-palatal [ʨ, ʨʰ, ɕ] or a new form with the noise pattern similar to those for the denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ].

Cantonese Female 4 (Table 13) is the only speaker in this study who can produce the distinct noise frequency patterns for the three types of Mandarin sibilants similar to those for the Mandarin speaker's sibilants (Table 5). As shown in Table 13 for Cantonese Female 4, the frequency values for both the noise peak and minimum value are the highest for the denti-alveolar [ʦ, ʦʰ, s] (9,383 Hz and 7,824 Hz), followed by the alveolo-palatal [ʨ, ʨʰ, ɕ] (6,754 Hz and 5,416 Hz) and then the retroflex [ʈʂ, ʈʂʰ, ʂ] (4,482 Hz and 2,782 Hz) in descending order. These frequency values are close to those for the corresponding sibilants [ʦ, ʦʰ, s] (9,145 Hz and 8,025 Hz), [ʨ, ʨʰ, ɕ] (6,967 Hz and 5,963 Hz) and [ʈʂ, ʈʂʰ, ʂ] (3,648 Hz and 2,059 Hz) of the Mandarin speaker as shown in Table 5.

To sum up, except for one speaker (Cantonese Female 4), all the other seven Cantonese speakers cannot produce the three distinct place categories of sibilants in Mandarin. For most of the speakers, the three types of sibilants may be produced as the denti-alveolar [ʦ, ʦʰ, s], alveolo-palatal [ʨ, ʨʰ, ɕ], or a 'new form' with the noise pattern similar to those for [ʦ, ʦʰ, s] and [ʨ, ʨʰ, ɕ]. Of the three place categories of Mandarin sibilants, the retroflex [ʈʂ, ʈʂʰ, ʂ] are the most problematic ones, merging to the other two types of Mandarin sibilants in the speech of Cantonese speakers.

3.1.2. Production of the Mandarin sibilants in Cantonese speakers

Tables 14-22 present the correct and incorrect production rates of the three place categories of Mandarin sibilants for each of the eight Cantonese speakers. The data are based on the comparison of the two acoustic attributes of the noise pattern, the noise peak and minimum value, for the sibilants produced by the Cantonese speakers and the native Mandarin speaker. In the table, the incorrect sibilants produced by Cantonese speakers are further indicated whether they become the other sibilants or a 'new form' which is unclassified to any type of sibilants in Mandarin.

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ts, tsʰ, s]	7/9 (77.78%)		0	0	2/9 (22.22%)
Retroflex [tʂ, tʂʰ, ʂ]	1/9 (11.11%)	1/9 (11.11%)		0	7/9 (77.78%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	0	2/9 (22.22%)	0		7/9 (77.78%)
Overall	**8/27 (29.63%)**	**3/27 (11.11%)**			**16/27 (59.26%)**

Table 14. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Male 1 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ts, tsʰ, s]	6/9 (66.67%)		0	0	3/9 (33.33%)
Retroflex [tʂ, tʂʰ, ʂ]	0	5/9 (55.56%)		0	4/9 (44.44%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	0	3/9 (33.33%)	0		6/9 (66.67%)
Overall	**6/27 (22.22%)**	**8/27 (29.63%)**			**13/27 (48.15%)**

Table 15. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Male 2 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	1/9 (11.11%)		0	0	8/9 (88.89%)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	1/9 (11.11%)	2/9 (22.22%)		0	6/9 (66.67%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	0	4/9 (44.44%)	0		5/9 (55.56%)
Overall	**2/27 (7.41%)**	**6/27 (22.22%)**			**19/27 (70.37%)**

Table 16. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Male 3 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	9/9 (100%)		0	0	0
Retroflex [ʈʂ, ʈʂʰ, ʂ]	5/9 (55.56%)	0		0	4/9 (44.44%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	1/9 (11.11%)	4/9 (44.44%)	0		4/9 (44.44%)
Overall	**15/27 (55.56%)**	**4/27 (14.81%)**			**8/27 (29.63%)**

Table 17. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Male 4 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	0		1/9 (11.11%)	1/9 (11.11%)	7/9 (77.78%)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	3/9 (33.33%)	0		1/9 (11.11%)	5/9 (55.56%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	4/9 (44.44%)	0	0		5/9 (55.56%)
Overall	**7/27 (25.93%)**	**3/27 (11.11%)**			**17/27 (62.96%)**

Table 18. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Female 1 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	0		1/9 (11.11%)	1/9 (11.11%)	7/9 (77.78%)
Retroflex [ʂ, ʂʰ, ʂ]	0	2/9 (22.22%)		1/9 (11.11%)	6/9 (66.67%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	1/9 (11.11%)	0	0		8/9 (88.89%)
Overall	**1/27 (3.70%)**	**5/27 (18.52%)**			**21/27 (77.78%)**

Table 19. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Female 2 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	5/9 (55.56%)		0	0	4/9 (44.44%)
Retroflex [ʂ, ʂʰ, ʂ]	0	3/9 (33.33%)		3/9 (33.33%)	3/9 (33.33%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	1/9 (11.11%)	1/9 (11.11%)	0		7/9 (77.78%)
Overall	**6/27 (22.22%)**	**7/27 (25.93%)**			**14/27 (51.85%)**

Table 20. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Female 3 (9 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	5/9 (55.56%)		0	0	4/9 (44.44%)
Retroflex [ʂ, ʂʰ, ʂ]	7/9 (77.78%)	0		0	2/9 (22.22%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	7/9 (77.78%)	0	0		2/9 (22.22%)
Overall	**19/27 (70.37%)**	**0**			**8/27 (29.63%)**

Table 21. Correct and incorrect rates of the production of Mandarin sibilants for Cantonese Female 4 (9 tokens for each place category; shaded area = impossible result).

A comparison of the production results for the eight Cantonese speakers presented in Tables 14-22 shows that Cantonese Female 4 has the highest overall correct rate (70.37%) for the production of the three types of Mandarin sibilants (Table 21). Her correct rate is especially high, approximately 80% for both the retroflex [tʂ, tʂʰ, ʂ] (77.78%) and alveolo-palatal [tɕ, tɕʰ, ɕ] (77.78%). The correct rate for the denti-alveolar [ts, tsʰ, s] is relatively low (55.56%), but it is still over 50% and larger than the incorrect rate (44.44%). In the incorrect cases, the three types of sibilants [ts, tsʰ, s], [tʂ, tʂʰ, ʂ] and [tɕ, tɕʰ, ɕ] are not mispronounced as another type of sibilants, but become a 'new form'. The data indicate that Cantonese Female 4 basically can clearly distinguish the three types of Mandarin sibilants.

As for the other seven Cantonese speakers, Cantonese Male 4 is the only one who has the overall correct rate over 50% as shown in Table 17. For this speaker, the correct rate largely varies among the three place categories of sibilants. It is 100% correct for the denti-alveolar [ts, tsʰ, s], but the correct rate reduces to 55.56% for the retroflex [tʂ, tʂʰ, ʂ] and it is only 11.11% for the alveolo-palatal [tɕ, tɕʰ, ɕ]. The data indicate that Cantonese Male 4 only has the problem in producing the retroflex [tʂ, tʂʰ, ʂ] and especially the alveolo-palatal [tɕ, tɕʰ, ɕ]. In the correct cases, the retroflex [tʂ, tʂʰ, ʂ] become a 'new form' only, but the alveolo-palatal [tɕ, tɕʰ, ɕ] may either be mispronounced as the denti-alveolar [ts, tsʰ, s] (44.44%) or become a 'new form' (44.44%). Thus, the main problem for Cantonese Male 4 is in the distinction between the alveolo-palatal [tɕ, tɕʰ, ɕ] and denti-alveolar [ts, tsʰ, s], where he mixed [tɕ, tɕʰ, ɕ] up with [ts, tsʰ, s].

For Cantonese Male 1 (Table 14) and 2 (Table 15) and Cantonese Female 1 (Table 18) and 3 (Table 20), their overall correct rates are in the range of 20-30% (i.e., 29.63%, 22.22%, 25.93%, and 22.22%, respectively). For Cantonese Male 1-2 and Cantonese Female 3, the correct rate is over 50% for the denti-alveolar [ts, tsʰ, s] (i.e.,

77.78%, 66.67%, and 55.56%, respectively). Their problem is in the production of the retroflex [tʂ, tʂʰ, ʂ] and alveolo-palatal [tɕ, tɕʰ, ɕ], where the correct rate is only 11.11% or 0%. For Cantonese Male 1, both the retroflex [tʂ, tʂʰ, ʂ] (77.78%) and alveolo-palatal [tɕ, tɕʰ, ɕ] (77.78%) become a 'new form' in most of the incorrect cases. For Cantonese Male 2, the retroflex [tʂ, tʂʰ, ʂ] and alveolo-palatal [tɕ, tɕʰ, ɕ] are mispronounced as the denti-alveolar [ts, tsʰ, s] in some incorrect cases (55.56% and 33.33%, respectively) and become a 'new form' in other incorrect cases (44.44% and 66.67%, respectively). For Female Cantonese 3, the alveolo-palatal [tɕ, tɕʰ, ɕ] become a 'new form' in most of the incorrect cases (77.78%), whereas the retroflex [tʂ, tʂʰ, ʂ] are mispronounced as the denti-alveolar [ts, tsʰ, s] (33.33%) and alveolo-palatal [tɕ, tɕʰ, ɕ] (33.33%), or become a 'new form' (33.33%) in the incorrect cases.

As for Cantonese Female 1, her production of the three types of Mandarin sibilants is problematic, as the correct rate is below 50% for [ts, tsʰ, s] (0%), [tʂ, tʂʰ, ʂ] (33.33%), and [tɕ, tɕʰ, ɕ] (44.44%). The worst set is the denti-alveolar [ts, tsʰ, s], which becomes a 'new form' in most of the incorrect cases (77.78%). The retroflex [tʂ, tʂʰ, ʂ] and alveolo-palatal [tɕ, tɕʰ, ɕ] also become a 'new form' in most of the incorrect cases (55.56% and 55.56%, respectively). Thus, the data may indicate that Cantonese Female 1 basically has no problem in distinguishing among the three types of Mandarin sibilants, but she cannot produce the sibilants similar to the native's ones.

The two remaining speakers, Cantonese Male 3 (Table 16) and Cantonese Female 2 (Table 19), performed badly in the production of the three types of Mandarin sibilants, with the overall correct rate below 10%. For Cantonese Male 3, the denti-alveolar [ts, tsʰ, s] only become a 'new form' in the incorrect cases (88.89%). As for the retroflex [tʂ, tʂʰ, ʂ] and alveolo-palatal [tɕ, tɕʰ, ɕ], they mainly become a 'new form' in most of the incorrect cases (66.67% and 55.56%, respectively), but may be mispronounced as the denti-alveolar [ts, tsʰ, s] in some incorrect cases (22.22% and

44.44%, respectively). For Cantonese Female 2, all the three types of sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] mainly become a 'new form' in the incorrect cases (77.78%, 66.67%, and 88.89%, respectively), though the denti-alveolar [ʦ, ʦʰ, s] and retroflex [ʈʂ, ʈʂʰ, ʂ] may be mispronounced as the other types of sibilants in few incorrect cases.

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	33/72 (45.83%)		2/72 (2.78%)	2/72 (2.78%)	35/72 (48.41%)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	17/72 (23.61%)	13/72 (18.06%)		5/72 (6.94%)	37/72 (51.39%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	14/72 (19.44%)	14/72 (19.44%)	0		44/72 (61.11%)
Overall	**64/216 (29.63%)**	**36/216 (16.67%)**			**116/216 (53.70%)**

Table 22. Correct and incorrect rates of the production of Mandarin sibilants for eight Cantonese speakers (72 tokens for each place category; shaded area = impossible result).

Table 22 presents the production results based on the data of all the eight Cantonese speakers. As shown in the table, the overall correct rate (29.63%) is much lower than the incorrect rate (70.37%). Among the three place categories of Mandarin sibilants, the performance is relatively better in the production of the denti-alveolar [ʦ, ʦʰ, s], with the correct rate of 45.83%, than [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ], with the correct rate of 23.61% and 19.44%, respectively. In the incorrect cases, the denti-alveolar [ʦ, ʦʰ, s] mainly become a 'new form' (48.41%). The retroflex [ʈʂ, ʈʂʰ, ʂ] and alveolo-palatal [ʨ, ʨʰ, ɕ] also become a 'new form' in most of the incorrect cases (51.39% and 61.11%, respectively), but they may be mispronounced as the denti-alveolar [ʦ, ʦʰ, s] in some other cases (18.06% and 19.44%, respectively). The

data in general may suggest that the Cantonese speakers are able to distinguish the three place categories of Mandarin sibilants, but they cannot produce the sibilants similar to the native sibilants in most cases.

3.1.3. Comparison of the noise patterns of Mandarin and Cantonese sibilants

As observed in all the eight Cantonese speakers, there are a number of Mandarin sibilants produced as a 'new form' with the noise pattern that cannot be classified to any one of the three place categories. In view of the fact that the only set of alveolar sibilants in Cantonese, i.e., [ʦ, ʦʰ, s], is different from the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ] in Mandarin, it is wondered whether the 'new form' of Mandarin sibilants in Cantonese speakers' L2 is derived under the influence of their L1 Cantonese. To answer this question, the frequency data of the noise patterns, in terms of both the noise peak and minimum value, of (i) the 'new forms' of Mandarin sibilants and (ii) the Cantonese sibilants [ʦ, ʦʰ, s] produced by each of the eight Cantonese speakers are compared. The results are presented in Table 23 for four male speakers and in Table 24 for four female speakers.

	Produced as a 'new form' similar to Cantonese [ʦ, ʦʰ, s]			
Target sibilants	**Cantonese Male 1**	**Cantonese Male 2**	**Cantonese Male 3**	**Cantonese Male 4**
Denti-alveolar [ʦ, ʦʰ, s]	0% (0/2)	0% (0/3)	16.67% (1/6)	0% (0)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	0% (0/7)	25% (1/4)	16.67% (1/6)	0% (0/4)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	14.29% (1/7)	0% (0/6)	20% (1/5)	0% (0/4)
Overall	**6.25% (1/16)**	**7.69% (1/13)**	**17.65% (3/17)**	**0% (0/8)**

Table 23. Percentages of the Mandarin sibilants produced as a 'new form' similar to the Cantonese sibilants [ʦ, ʦʰ, s] for Cantonese Male 1-4 (no. of tokens out of the total in parentheses).

Target sibilants	Produced as a 'new form' similar to Cantonese [ʦ, ʦʰ, s]			
	Cantonese Female 1	**Cantonese Female 2**	**Cantonese Female 3**	**Cantonese Female 4**
Denti-alveolar [ʦ, ʦʰ, s]	42.86% (3/7)	28.57% (2/7)	0% (0/4)	0% (0)
Retroflex [ʧ, ʧʰ, ʃ]	0% (0/5)	16.67% (1/6)	0% (0/3)	0% (0)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	20% (1/5)	50% (4/8)	14.29% (1/7)	0% (0)
Overall	**23.53%(4/17)**	**33.33%(7/21)**	**7.14% (1/14)**	**0% (0)**

Table 24. Percentages of the Mandarin sibilants produced as a 'new form' similar to the Cantonese sibilants [ʦ, ʦʰ, s] for Cantonese Female 1-4 (no. of tokens out of the total in parentheses).

For the four male Cantonese speakers (Table 23), it can be seen that the overall percentage of Mandarin sibilants that are produced as a 'new form' and similar to the Cantonese [ʦ, ʦʰ, s] is noticeably low, in the range of 0-17.65%. With regard to the 'new form' cases for each one of the three place categories of Mandarin sibilants, the percentage is also low, which is in the range of 0-25% for the four male Cantonese speakers.

As for the female Cantonese speakers (Table 24), excluding Cantonese Female 4 who have no Mandarin sibilants produced as a 'new form', the overall percentage for the 'new forms' which are similar to the Cantonese [ʦ, ʦʰ, s] is also low, in the range of 7.14-33.33%. However, considering the 'new form' cases for each of the three types of Mandarin sibilants, it is observed that for Cantonese Female 1, there are 42.86% of the 'new form' cases for the Mandarin [ʦ, ʦʰ, s] similar to the Cantonese [ʦ, ʦʰ, s]. For Cantonese Female 2, there are 50% of the 'new form' cases for the Mandarin [ʨ, ʨʰ, ɕ] similar to the Cantonese [ʦ, ʦʰ, s]. Thus, it may be considered that there is a tendency for these two female speakers to use the Cantonese [ʦ, ʦʰ, s] to produce the denti-alveolar [ʦ, ʦʰ, s] or alveolo-palatal [ʨ, ʨʰ, ɕ] in

Mandarin. But, in general, the influence of L1 of the Cantonese speakers on the production of Mandarin sibilants in L2 is not significant.

Table 25 presents the percentages of the 'new form' Mandarin sibilants that are similar to the Cantonese [ʦ, ʦʰ, s] for all the eight Cantonese speakers. Some observations are made as follows. (i) The overall percentage of the 'new form' cases in which the Mandarin sibilants are similar to the Cantonese [ʦ, ʦʰ, s] is only 16.04%. (ii) Such kind of percentage is slightly increased with respect to the 'new forms' for the Mandarin [ʦ, ʦʰ, s] (20.69%) or [ʨ, ʨʰ, ɕ] (19.05%), and it is largely decreased for the Mandarin [ʈʂ, ʈʂʰ, ʂ] (8.57%). (iii) In general, no significant influence of L1 Cantonese on the production of Mandarin sibilants in L2 is observed, although the L1 effect seems to be slightly increased on the production of the Mandarin denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ] which are similar to the Cantonese alveolar [ʦ, ʦʰ, s] in articulation.

Target sibilants	**Produced as a 'new form' similar to Cantonese [ʦ, ʦʰ, s]**
Denti-alveolar [ʦ, ʦʰ, s]	20.69% (6/29)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	8.57% (3/35)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	19.05% (8/42)
Overall	**16.04% (17/106)**

Table 25. Percentages of the Mandarin sibilants produced as a 'new form' similar to the Cantonese sibilants [ʦ, ʦʰ, s] for eight Cantonese speakers (no. of tokens out of the total in parentheses).

3.1.4. Perceptual assessment of the Mandarin sibilants for Cantonese speakers

The production of the Mandarin sibilants from the eight Cantonese speakers was also perceptually assessed by a group of 10 Mandarin native speakers. The listeners were instructed to judge whether the Mandarin sibilants produced by the

Cantonese speakers are ‘correct’ or ‘incorrect’, and for the ‘incorrect cases’ whether the sibilants are mispronounced as the other type of sibilants or ‘NA’, i.e., unclassifiable to any type of Mandarin sibilants. Tables 26-33 present the results of perceptual assessment for the Mandarin sibilants from each of the eight Cantonese speakers.

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	56/90 (62.22%)		31/90 (34.45%)	0	3/90 (3.33%)
Retroflex [tʂ, tʂʰ, ʂ]	71/90 (78.89%)	16/90 (17.78%)		0	3/90 (3.33%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	73/90 (81.11%)	3/90 (3.33%)	1/90 (1.11%)		13/90 (14.44%)
Overall	**200/270 (74.07%)**	**51/270 (18.89%)**			**19/270 (7.04%)**

Table 26. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Male 1 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	82/90 (91.11%)		5/90 (5.56%)	0	3/90 (3.33%)
Retroflex [tʂ, tʂʰ, ʂ]	68/90 (75.56%)	19/90 (21.11%)		0	3/90 (3.33%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	80/90 (88.89%)	0	2/90 (2.22%)		8/90 (8.89%)
Overall	**230/270 (85.19%)**	**26/270 (9.63%)**			**14/270 (5.18%)**

Table 27. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Male 2 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	9/90 (10%)		69/90 (76.67%)	0	12/90 (13.33%)
Retroflex [tʂ, tʂʰ, ʂ]	76/90 (84.44%)	14/90 (15.56%)		0	0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	72/90 (80%)	0	0		18/90 (20%)
Overall	**157/270 (58.15%)**	**83/270 (30.74%)**			**30/270 (11.11%)**

Table 28. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Male 3 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	90/90 (100%)		0	0	0
Retroflex [tʂ, tʂʰ, ʂ]	88/90 (97.78%)	2/90 (2.22%)		0	0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	67/90 (74.44%)	1/90 (1.11%)	0		22/90 (24.44%)
Overall	**245/270 (90.74%)**	**3/270 (1.11%)**			**22/270 (8.15%)**

Table 29. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Male 4 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	56/90 (62.22%)		30/90 (33.33%)	0	4/90 (4.44%)
Retroflex [tʂ, tʂʰ, ʂ]	49/90 (54.44%)	14/90 (15.56%)		10/90 (11.11%)	17/90 (18.89%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	58/90 (64.44%)	22/90 (24.44%)	9/90 (10%)		1/90 (1.11%)
Overall	**163/270 (60.37%)**	**85/270 (31.48%)**			**22/270 (8.15%)**

Table 30. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Female 1 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	13/90 (14.44%)		74/90 (82.22%)	0	3/90 (3.33%)
Retroflex [tʂ, tʂʰ, ʂ]	74/90 (82.22%)	11/90 (12.22%)		0	5/90 (5.56%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	90/90 (100%)	0	0		0
Overall	**177/270 (65.56%)**	**85/270 (31.48%)**			**8/270 (2.96%)**

Table 31. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Female 2 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	63/90 (70%)		6/90 (6.67%)	0	21/90 (23.33%)
Retroflex [tʂ, tʂʰ, ʂ]	72/90 (80%)	15/90 (16.67%)		0	3/90 (3.33%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	82/90 (91.11%)	1/90 (1.11%)	0		7/90 (7.78%)
Overall	**217/270 (80.37%)**	**22/270 (8.15%)**			**31/270 (11.48%)**

Table 32. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Female 3 (90 responses for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	NA
Denti-alveolar [ts, tsʰ, s]	88/90 (97.78%)		2/90 (2.22%)	0	0
Retroflex [tʂ, tʂʰ, ʂ]	89/90 (98.89%)	1/90 (1.11%)		0	0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	90/90 (100%)	0	0		0
Overall	**267/270 (98.89%)**	**3/270 (1.11%)**			**0**

Table 33. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for Cantonese Female 4 (90 responses for each place category; shaded area = impossible result).

Based on the perceptual judgement of native Mandarin speakers, the overall correct production rate of the Mandarin sibilants is markedly increased and over 50% for all the eight Cantonese speakers as shown in Tables 26-33. The highest correct rate is for Cantonese Female 4, which is near 100% (Table 33). A comparison of the production results for Cantonese Female 4 based on the perceptual judgement (Table 33) and the frequency data of the sibilants (Table 21) shows that most of the tokens classified as a 'new form' based on the frequency data are judged as 'correct' based on the Mandarin speakers' perception. Based on the frequency data, the rate of Cantonese Female 4's sibilants classified as a 'new form' is 29.63% versus the correct rate of 70.37%. Based on the perceptual judgement, Cantonese Female 4's sibilants judged as 'NA' is 0% versus the correct rate of 98.89%. The difference in the correct rate between the two sets of data may suggest that the deviation of the 'new form' from the 'target' in terms of the noise frequency is not large and the 'new form' is still within the categorical boundary of the 'target' in the perception of Mandarin speakers.

Similar case of the increase in the correct rate and the corresponding drop in the rate of 'new form/NA' based on the perceptual judgement is also observed in all the other seven Cantonese speakers. Based on the frequency data, the overall correct rates and rates of 'new form' are 29.63% and 59.26% (Cantonese Male 1; Table 14), 22.22% and 48.15% (Cantonese Male 2; Table 15), 7.41% and 70.37% (Cantonese Male 3; Table 16), 55.56% and 29.63% (Cantonese Male 4; Table 17), 25.93% and 62.96% (Cantonese Female 1; Table 18), 3.70% and 77.78% (Cantonese Female 2; Table 19), and 22.22% and 51.85% (Cantonese Female 3; Table 20). Based on the perceptual judgement, the overall correct rates and rates of 'NA' are 74.07% and 7.04% (Cantonese Male 1; Table 26), 85.19% and 5.18% (Cantonese Male 2; Table 27), 58.15% and 11.11% (Cantonese Male 3; Table 28), 90.74% and 8.15% (Cantonese Male 4; Table 29), 60.37% and 8.15% (Cantonese Female 1; Table 30),

65.56% and 2.96% (Cantonese Female 2; Table 31), and 80.37% and 11.48% (Cantonese Female 3; Table 32).

Based on the perceptual judgement, a high overall correct production rate over 80% is observed for Cantonese Male 2 (85.19%), Cantonese Male 4 (90.74%), and Cantonese Female 3 (80.37%), in addition to Cantonese Female 4 (98.89%). For these speakers, the correct production rate is over 70% for each one of the three place categories of Mandarin sibilants, which indicates that the pronunciation of their Mandarin sibilants is basically identifiable and acceptable to the native speakers.

As for the other Cantonese speakers (Cantonese Male 1 & 3 and Cantonese Female 1 & 2), based on the perceptual judgement, there is a decrease in their overall correct production rates (in the range of 58-74%), with a corresponding increase in their overall incorrect production rates (in the range of 26-42%). Their major problem is the production of the denti-alveolar [ʦ, ʦʰ, s], with 34.45% of [ʦ, ʦʰ, s] mispronounced as the retroflex [ʧ, ʧʰ, ʂ] for Cantonese Male 1, 76.67% of [ʦ, ʦʰ, s] → [ʧ, ʧʰ, ʂ] for Cantonese Male 3, 33.33% of [ʦ, ʦʰ, s] → [ʧ, ʧʰ, ʂ] for Cantonese Female 1, and 82.22% of [ʦ, ʦʰ, s] → [ʧ, ʧʰ, ʂ] for Cantonese Female 2. For these speakers, the number of cases for [ʧ, ʧʰ, ʂ] mispronounced as [ʦ, ʦʰ, s] is relatively small, i.e., 17.78% for Cantonese Male 1, 15.56% for Cantonese Male 3, 15.56% for Cantonese Female 1, and 12.22% for Cantonese Female 2. This indicates that the Cantonese speakers mainly mix [ʦ, ʦʰ, s] with [ʧ, ʧʰ, ʂ], but not in the other way round. For Cantonese Female 1, she also has problem in the production of the alveolo-palatal [ʨ, ʨʰ, ɕ], where some of the alveolo-palatal [ʨ, ʨʰ, ɕ] from her are perceptually judged as the denti-alveolar [ʦ, ʦʰ, s] (24.44%) by Mandarin speakers. All these data indicate that it is more problematic for Cantonese speakers to produce the denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ] in Mandarin than the

retroflex [ʦ, ʦʰ, s], where the former two sets of Mandarin sibilants but not the latter set are similar to the Cantonese alveolar [ʦ, ʦʰ, s].

The results of perceptual judgement of the Mandarin sibilants produced by all the eight Cantonese speakers are presented in Table 34. The data in the table show that the three place categories of Mandarin sibilants are basically identifiable in the speech of Cantonese speakers, with the overall correct production rate of 76.67%. Among the three place categories, the correct rate is relatively smaller for the denti-alveolar [ʦ, ʦʰ, s] (63.47%) than the retroflex [ʈʂ, ʈʂʰ, ʂ] (81.53%) and alveolo-palatal [ʨ, ʨʰ, ɕ] (85%), indicating that the difficulty for Cantonese speakers is more in the production of the denti-alveolar [ʦ, ʦʰ, s] than the other two types of sibilants in Mandarin. For the incorrect cases, most of them are the mispronunciation cases (with the overall rate of 16.57%), rather than the NA cases (with the overall rate of 6.76%). The mispronunciation is mainly for the denti-alveolar [ʦ, ʦʰ, s] to become the retroflex [ʈʂ, ʈʂʰ, ʂ] (30.14%), while [ʈʂ, ʈʂʰ, ʂ] are also mispronounced as [ʦ, ʦʰ, s] in some cases (12.78%).

Target sibilants	**Correct**	**Incorrect**			
		Denti-alveolar	**Retroflex**	**Alveolo-palatal**	**NA**
Denti-alveolar [ʦ, ʦʰ, s]	457/720 (63.47%)		217/720 (30.14%)	0	46/720 (6.39%)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	587/720 (81.53%)	92/720 (12.78%)		10/720 (1.39%)	31/720 (4.30%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	612/720 (85%)	27/720 (3.75%)	12/720 (1.67%)		69/720 (9.58%)
Overall	**1656/2160 (76.67%)**	**358/2160 (16.57%)**			**146/2160 (6.76%)**

Table 34. Correct and incorrect rates of the Mandarin sibilants perceptually assessed by ten Mandarin speakers for eight Cantonese speakers (720 responses for each place category; shaded area = impossible result).

3.2. Results of Perception Experiment

In addition to production, the Cantonese speakers of this study participated in a perception experiment to identify the three place categories of Mandarin sibilants produced by a native Mandarin female speaker. Tables 35-42 present the identification rates of the different types of Mandarin sibilants for each of the eight Cantonese speakers.

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	11/15 (73.33%)		4/15 (26.67%)	0
Retroflex [tʂ, tʂʰ, ʂ]	12/15 (80.00%)	3/15 (20.00%)		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	15/15 (100%)	0	0	
Overall	**38/45 (84.44%)**	**7/45 (15.56%)**		

Table 35. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Male 1 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	14/15 (93.33%)		1/15 (6.67%)	0
Retroflex [tʂ, tʂʰ, ʂ]	15/15 (100%)	0		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	15/15 (100%)	0	0	
Overall	**44/45 (97.78%)**	**1/45 (2.22%)**		

Table 36. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Male 2 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	9/15 (60.00%)		5/15 (33.33%)	1/15 (6.67%)
Retroflex [tʂ, tʂʰ, ʂ]	11/15 (73.33%)	2/15 (13.33%)		2/15 (13.33%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	10/15 (66.67%)	4/15 (26.67%)	1/15 (6.67%)	
Overall	**30/45 (66.67%)**	**15/45 (33.33%)**		

Table 37. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Male 3 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	15/15 (100%)		0	0
Retroflex [tʂ, tʂʰ, ʂ]	15/15 (100%)	0		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	15/15 (100%)	0	0	
Overall	**45/45 (100%)**	**0**		

Table 38. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Male 4 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	14/15 (93.33%)		0	1/15 (6.67%)
Retroflex [tʂ, tʂʰ, ʂ]	15/15 (100%)	0		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	13/15 (86.67%)	1/15 (6.67%)	1/15 (6.67%)	
Overall	**42/45 (93.33%)**	**3/45 (6.67%)**		

Table 39. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Female 1 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	13/15 (86.67%)		1/15 (6.67%)	1/15 (6.67%)
Retroflex [tʂ, tʂʰ, ʂ]	15/15 (100%)	0		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	15/15 (100%)	0	0	
Overall	**43/45 (95.56%)**	**2/45 (4.44%)**		

Table 40. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Female 2 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	13/15 (86.67%)		2/15 (13.33%)	0
Retroflex [tʂ, tʂʰ, ʂ]	15/15 (100%)	0		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	13/15 (86.67%)	2/15 (13.33%)	0	
Overall	**41/45 (91.11%)**	**4/45 (8.89%)**		

Table 41. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Female 3 (15 stimuli for each place category; shaded area = impossible result).

Target stimuli	Correct	Incorrect		
		Denti-alveolar	Retroflex	Alveolo-palatal
Denti-alveolar [ts, tsʰ, s]	14/15 (93.33%)		1/15 (6.67%)	0
Retroflex [tʂ, tʂʰ, ʂ]	15/15 (100%)	0		0
Alveolo-palatal [tɕ, tɕʰ, ɕ]	15/15 (100%)	0	0	
Overall	**44/45 (97.78%)**	**1/45 (2.22%)**		

Table 42. Correct and incorrect identification rates of the Mandarin sibilants for Cantonese Female 4 (15 stimuli for each place category; shaded area = impossible result).

In the perception experiment, most of the Cantonese speakers (Cantonese Male 2 & 4 and Female Cantonese 1-4), except two (Cantonese Male 1 & 3), performed very well in the identification of the Mandarin sibilants, with an overall correct identification rate over 90%. Cantonese Male 4 performed the best with 100% correct identification rate (Table 38). Cantonese Male 2 (Table 36) and Cantonese Female 4 (Table 42) also correctly identified all the sibilants, but one in the denti-alveolar group, with the overall correct rate of 97.78%. For Cantonese Female 2 (Table 40), the overall correct identification rate is 95.56%, with two wrong cases also for the denti-alveolar sibilants. As for Cantonese Female 1 (Table 39) and Female 3 (Table 41), the overall correct identification rates are 93.33% and 91.11%. For these two speakers, there are 1-2 wrong identification cases for the denti-alveolar [ʦ, ʦʰ, s] and alveolo-palatal [ʨ, ʨʰ, ɕ] sibilants, but not the retroflex sibilants [ʈʂ, ʈʂʰ, ʂ].

As for the two remaining speakers, Cantonese Male 1 (Table 35) performed better, with the overall correct identification rate of 84.44%, than Cantonese Male 3 (Table 37), with the overall correct identification rate of 66.67%. For Cantonese Male 1, he identified correctly for all the alveolo-palatal [ʨ, ʨʰ, ɕ], but wrongly identified the denti-alveolar [ʦ, ʦʰ, s] as [ʈʂ, ʈʂʰ, ʂ] (26.67%) and the retroflex [ʈʂ, ʈʂʰ, ʂ] as [ʦ, ʦʰ, s] (20%) in some cases. For Cantonese Male 3, all the three place categories of Mandarin sibilants are wrongly identified as the other types of sibilants, with a slightly higher wrong identification rate for the denti-alveolar [ʦ, ʦʰ, s] (40%), followed by the alveolo-palatal [ʨ, ʨʰ, ɕ] (33.33%) and then the retroflex [ʈʂ, ʈʂʰ, ʂ] (26.67%) in descending order.

Overall, based on the perception results of all the eight Cantonese speakers as presented in Table 43, Cantonese speakers have a high level of competence in the identification of the three place categories of Mandarin sibilants, with a large correct rate of 90.83% and a minimal incorrect rate of 9.17%. Among the three place

categories, the correct rate is over 90% for both the retroflex [tʂ, tʂʰ, ʂ] (94.17%) and alveolo-palatal [tɕ, tɕʰ, ɕ] (92.5%) and it is slightly lowered to 85.83% for the denti-alveolar [ts, tsʰ, s]. In the incorrect cases, all the three types of Mandarin sibilants may be misidentified as the other two types of sibilants, while the denti-alveolar [ts, tsʰ, s] are more frequently misidentified as the retroflex [tʂ, tʂʰ, ʂ]. In general, the perceptual ability of the Cantonese speakers is slightly weaker to identify [ts, tsʰ, s] than [tʂ, tʂʰ, ʂ] and [tɕ, tɕʰ, ɕ].

Target stimuli	**Correct**	**Incorrect**		
		Denti-alveolar	**Retroflex**	**Alveolo-palatal**
Denti-alveolar [ts, tsʰ, s]	103/120 (85.83%)		14/120 (11.67%)	3/120 (2.5%)
Retroflex [tʂ, tʂʰ, ʂ]	113/120 (94.17%)	5/120 (4.17%)		2/120 (1.66%)
Alveolo-palatal [tɕ, tɕʰ, ɕ]	111/120 (92.5%)	7/120 (5.83%)	2/120 (1.66%)	
Overall	**327/360 (90.83%)**	**33/360 (9.17%)**		

Table 43. Correct and incorrect identification rates of the Mandarin sibilants for eight Cantonese speakers (120 stimuli for each place category; shaded area = impossible result).

4. Discussion

In this section, the results of the production and perception experiments of Mandarin sibilants for the Cantonese speakers presented in the present study are to be used (i) to compare with those reported in the previous studies of the production (Wong, 2015) and perception (Lai 2009) of Mandarin sibilants, (ii) to determine (a) the similarities and differences between the phonetic realization in production and phonological categorization in perception for the L2 sounds and (b) the performance of production and perception in L2 learners, (iii) to discuss the L1 influence on L2 production and perception, and (iv) to evaluate the prediction of the speech learning model of L2 acquisition proposed by Flege (1995).

4.1. Comparison with the Previous Studies

4.1.1. Wong (2015)

In Wong (2015), the noise peak and noise range of the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ] produced by four Cantonese speakers, two male and two female, were analyzed and compared with those for a native Mandarin speaker. The correct and incorrect rates of the Mandarin sibilants production for the four Cantonese speakers in Wong (2015) are presented in Table 44 and used to compare with the corresponding data for the eight Cantonese speakers in the present study which are presented earlier in Table 22 and re-presented in Table 45 for easy comparison.

A comparison of Table 44 (Wong, 2015) and Table 45 (the present study) shows that the overall correct and incorrect production rates of the Mandarin sibilants are similar between the two studies. In both studies, the overall correct rate is below 50%, although the rate is slightly larger in Wong (2015) (35.19%) than the present

study (29.63%). The data indicate that the Cantonese speakers of both studies in general have not mastered the production of the different types of Mandarin sibilants.

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	8/72 (11.11%)		9/72 (12.5%)	19/72 (26.39%)	36/72 (50%)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	50/72 (69.44%)	7/72 (9.72%)		3/72 (4.17%)	12/72 (16.67%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	18/72 (25%)	1/72 (1.39%)	8/72 (11.11%)		45/72 (62.5%)
Overall	**76/216 (35.19%)**	**47/216 (21.76%)**			**93/216 (43.05%)**

Table 44. Correct and incorrect rates of the production of Mandarin sibilants for four Cantonese speakers in Wong (2015) (72 tokens for each place category; shaded area = impossible result).

Target sibilants	Correct	Incorrect			
		Denti-alveolar	Retroflex	Alveolo-palatal	New form
Denti-alveolar [ʦ, ʦʰ, s]	33/72 (45.83%)		2/72 (2.78%)	2/72 (2.78%)	35/72 (48.41%)
Retroflex [ʈʂ, ʈʂʰ, ʂ]	17/72 (23.61%)	13/72 (18.06%)		5/72 (6.94%)	37/72 (51.39%)
Alveolo-palatal [ʨ, ʨʰ, ɕ]	14/72 (19.44%)	14/72 (19.44%)	0		44/72 (61.11%)
Overall	**64/216 (29.63%)**	**36/216 (16.67%)**			**116/216 (53.70%)**

Table 45. Correct and incorrect rates of the production of Mandarin sibilants for eight Cantonese speakers in the present study (72 tokens for each place category; shaded area = impossible results).

Among the three place categories of Mandarin sibilants, the correct production rate is over 50% for the retroflex [ʈʂ, ʈʂʰ, ʂ] (69.44%) and much larger than

the correct rates of 11.11% for the denti-alveolar [ʦ, ʦʰ, s] and 25% for the alveolo-palatal [ʨ, ʨʰ, ɕ] in Wong (2015). In the present study, the correct production rate is below 50% for all the three place categories, while it is relatively larger for [ʦ, ʦʰ, s] (45.83%), followed by [ʈʂ, ʈʂʰ, ʂ] (23.61%) and then [ʨ, ʨʰ, ɕ] (19.44%) in descending order. The differences between the two studies denote the variation between Cantonese speakers in L2 production of Mandarin sibilants.

Considering the incorrect production cases, Cantonese speakers commonly produce the Mandarin sibilants to become a 'new form' which is unclassified to any one of the three place categories, with the overall rates of 43.05% in Wong (2015) and 53.07% in the present study, rather than mispronounce the Mandarin sibilants as the other categories, with the overall rates of 21.76% (Wong, 2015) and 16.67% (the present study). The data may indicate that Cantonese speakers are basically able to distinguish the three types of Mandarin sibilants, although the sibilants in their speech are not native-like.

4.1.2. Lai (2009)

Lai (2009) carried out a perception test of the ability of Malay and Burmese speakers to discriminate the three place categories of Mandarin affricates [ʦ, ʦʰ], [ʈʂ, ʈʂʰ], and [ʨ, ʨʰ]. The performance of the two groups of speakers in Lai (2009) is much worse than that of the Cantonese speakers in the present study who obtained the overall correct identification rate of the three types of Mandarin sibilants over 90% (Table 43).

Among the three different place categories of Mandarin sibilants, both the Malay and Burmese speakers in Lai (2009) performed better in the discrimination between the denti-alveolar [ʦ, ʦʰ] and alveolo-palatal [ʨ, ʨʰ] and between the retroflex [ʈʂ, ʈʂʰ] and alveolo-palatal [ʨ, ʨʰ] than between the denti-alveolar [ʦ, ʦʰ]

and retroflex [tʂ, tʂʰ]. As for the Cantonese speakers in the present study, their weakness is mainly in the discrimination between the denti-alveolar [ʦ, ʦʰ, s] and retroflex [tʂ, tʂʰ, ʂ]. The data demonstrate some similar error patterns in the perception of Mandarin sibilants between the subjects of different L1 in the two studies.

4.2. Relationship between Production and Perception of L2 Sounds

4.2.1. Phonetic realization and phonological/perceptual categorization

In the present study, the Mandarin sibilants produced by the Cantonese speakers are assessed based on the measured the frequency data on the noise patterns of the sibilants and also the perceptual judgement of a group of native Mandarin speakers. Based on the frequency data, the overall correct production rate of the Mandarin sibilants is below 50% for the Cantonese speakers, and the sibilants are frequently (with an overall rate of 53.7%) produced as a 'new form' which is unclassified to any one of the three place categories (Table 22). As for the production results based on the perceptual judgement of native Mandarin speakers, the overall correct production rate of the Mandarin sibilants from the Cantonese speakers is markedly increased to 76.67%, whereas the overall rate of the sibilants identified as a 'new form' is reduced to 6.76% (Table 34). The production results provided by the two different assessment methods demonstrate the discrepancy between the phonetic realization in production and phonological categorization in perception for the L2 Mandarin sibilant sounds. Most of the sibilants classified as a 'new form' based on the frequency data are not native-like, but they are still perceptually identifiable and acceptable for native Mandarin speakers, which indicates that the 'new form' different from the 'target' sibilant phonetically is still within the categorical or perceptual boundary of the 'target' sibilant phonologically.

Regarding the Mandarin sibilants that are mispronounced as the other types of sibilants by Cantonese speakers, the number of cases is low, with a rate of 16.67% based on the frequency data analysis (Table 22) and a rate of 16.57% based on the perceptual judgement (Table 34). A low percentage and a striking similarity for the two assessment methods confirm that the Cantonese speakers basically have no difficulty in distinguishing the three place categories of Mandarin sibilants, although the sibilants in their speech are not native-like in production as well as in perception.

4.2.2. Production and perception of L2 sounds

In the present study, both the data on the production and perception of the Mandarin sibilants were obtained from Cantonese speakers. A comparison of the two sets of data shows that the performance of Cantonese speakers is much better in perception, with a large overall correct rate of 90.83% for the identification of the Mandarin sibilants (Table 43), than in production, with a low overall correct production rate of 29.63% based on the frequency data analysis (Table 22) or a rate of 76.67% based on the perceptual assessment of native Mandarin speakers (Table 34). The result indicates that the competence in production and perception of L2 sounds may not be the same, and learners are easier to identify rather than to produce the L2 sounds.

Among the three place categories of Mandarin sibilants, Cantonese speakers are more frequently misidentified the denti-alveolar [ʦ, ʦʰ, s] (14.17%) than the retroflex [ʈʂ, ʈʂʰ, ʂ] (5.83%) and alveolo-palatal [ʨ, ʨʰ, ɕ] (7.5%) in perception (Table 43). Similarly, in production based on the perceptual judgement of native Mandarin speakers (Table 34), Cantonese speakers are also more frequently mispronounced the denti-alveolar [ʦ, ʦʰ, s] (30.14%) than the retroflex [ʈʂ, ʈʂʰ, ʂ] (14.17%) and alveolo-palatal [ʨ, ʨʰ, ɕ] (5.42%). The data demonstrate a similarity in production and

perception of the sibilants for Mandarin learners, with the denti-alveolar [ʦ, ʦʰ, s] as the most difficult type for Cantonese speakers.

4.3. L1 influence on L2 Production and Perception

In Cantonese, there is only one set of sibilants, the alveolar [ʦ, ʦʰ, s], corresponding to the three sets of sibilants, the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], in Mandarin. Thus, the L1 Cantonese influence on acquisition of L2 Mandarin, in particular the negative transfer of the Cantonese [ʦ, ʦʰ, s] to replace the three sets of Mandarin sibilants in production and the confusion of the three categories of Mandarin sibilants in perception are expected. Such expectation however is not supported by the production and perception data obtained in the present study.

In the production experiment of the present study, it is observed that the mispronunciation of the three place categories of Mandarin sibilants as the other categories is not common in the speech of Cantonese speakers. Instead, in most of the incorrect production cases, the Mandarin sibilants produced by Cantonese speakers become a ‘new form’ which is different from any one of the three place categories of sibilants in Mandarin. Furthermore, a comparison of the noise patterns for the Mandarin sibilants and Cantonese sibilants produced by Cantonese speakers shows that the Mandarin sibilants classified as a ‘new form’ are different from the Cantonese sibilants [ʦ, ʦʰ, s] in the speech of Cantonese speakers (see Table 25). It follows that no substitution of the Mandarin sibilants in L2 with the Cantonese sibilants in L1 is made by Cantonese speakers in the present study. The production results of the present study are not in agreement with the observations reported in the previous studies that (i) Cantonese speakers use the Cantonese [ʦ, ʦʰ, s] to replace the Mandarin [ʨ, ʨʰ, ɕ] (Ng, 2001) or to replace the Mandarin [ʦ, ʦʰ, s] and [ʈʂ, ʈʂʰ, ʂ]

(Hon, 2003), and (ii) Cantonese speakers mispronounce the Mandarin [ʦ, ʦʰ, s] and [ʧ, ʧʰ, ʃ] as [ʨ, ʨʰ, ɕ] (Lee-Wong, 2013).

In the perception experiment of the present study, the confusion of the three place categories of Mandarin sibilants is also not observed for Cantonese speakers. The overall correct rate is high in the range of 85-95 % for Cantonese speakers in the identification of all the three types or each one of the three types of Mandarin sibilants (Table 43). The perception data indicate that Cantonese speakers have no difficulty in distinguishing the three sets of sibilants [ʦ, ʦʰ, s], [ʧ, ʧʰ, ʃ], and [ʨ, ʨʰ, ɕ] in Mandarin, although there is only one set of alveolar sibilants [ʦ, ʦʰ, s] in Cantonese.

4.4. Flege's (1995) Speech Learning Model of L2 Acquisition

According to the Speech Learning Model of L2 acquisition proposed by Flege (1995), it is predicted that the 'same' phones in both L1 and L2 and the 'new' phones in L2 are easy to be acquired by learners, rather than the 'similar' phones between L1 and L2. In Cantonese, the only set of sibilants is the alveolar [ʦ, ʦʰ, s] made with the articulatory contact extending from the alveolar area to the postalveolar region (Zee, 1999). In Mandarin, although the sibilants [ʦ, ʦʰ, s] are represented with the alveolar symbols, they are the denti-alveolar sounds (Lee and Zee, 2003) and then they are just 'similar' to, but not the 'same' as the Cantonese [ʦ, ʦʰ, s]. As for the alveolo-palatal [ʨ, ʨʰ, ɕ] and retroflex [ʧ, ʧʰ, ʃ] in Mandarin, both of them are not occurring in Cantonese and then they are considered as the 'new' phones in L2 Mandarin. But, due to the fact that the Mandarin [ʨ, ʨʰ, ɕ] are produced with the articulatory contact extending from the pre-palatal area to the postalveolar area (Lee and Zee, 2003) more 'similar' to the Cantonese [ʦ, ʦʰ, s] than the Mandarin [ʧ, ʧʰ, ʃ] that are produced by retracting the tongue tip backward to the postalveolar area (Lee and Zee, 20013), the Mandarin [ʨ, ʨʰ, ɕ] are considered as the 'new similar phones'

and the Mandarin [tʂ, tʂʰ, ʂ] as the ‘new non-similar phones’. Therefore, under the Speech Learning Model, in L2 Mandarin acquisition for Cantonese speakers, the retroflex [tʂ, tʂʰ, ʂ] are the most easiest sounds, followed by the alveolo-palatal [tɕ, tɕʰ, ɕ] and then the denti-alveolar [ts, tsʰ, s] in descending order. Such order is supported by the perception data, but not the production data obtained in the present study.

In the perception experiment, the descending order of the overall correct identification rate is [tʂ, tʂʰ, ʂ] (94.17%) > [tɕ, tɕʰ, ɕ] (92.5%) > [ts, tsʰ, s] (85.83%) for Cantonese speakers (Table 43). However, in the production experiment, the order of the overall correct production rate based on the frequency data analysis (Table 22) is [ts, tsʰ, s] (45.83%) > [tʂ, tʂʰ, ʂ] (23.61%) > [tɕ, tɕʰ, ɕ] (19.44%). Based on the perceptual assessment of native Mandarin speakers (Table 34), the overall correct production rate is [tɕ, tɕʰ, ɕ] (85%) > [tʂ, tʂʰ, ʂ] (81.53%) > [ts, tsʰ, s] (63.47%). The production results in the present study are predicted by the Speech Learning Model (Flege, 1995), while the perception results give support to the model.

5. Conclusion

The present study investigates both the production and perception of the three place categories of Mandarin sibilants, namely the denti-alveolar [ʦ, ʦʰ, s], retroflex [ʈʂ, ʈʂʰ, ʂ], and alveolo-palatal [ʨ, ʨʰ, ɕ], by Cantonese-speaking university students. The production data on the Mandarin sibilants from Cantonese speakers were analyzed through performing acoustic analysis of the frequency of the noise patterns of the sibilants and carrying out perceptual assessment based on native Mandarin speakers' impression. The production results provided by the two methods show the Cantonese speakers basically have no difficulty in distinguishing the three different categories of Mandarin sibilants in production, while most of the Mandarin sibilants produced by Cantonese speakers are classified as a 'new form' different from any one of the three place categories of sibilants in Mandarin. Phonetically, the 'new form' sibilants are different from the native ones, but phonologically, they are still identifiable and acceptable within the categorical boundary of the 'target' Mandarin sibilants in perception. In general, the expectation that the negative L1 transfer of the Cantonese [ʦ, ʦʰ, s] to replace the three types of Mandarin sibilants, [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ], and [ʨ, ʨʰ, ɕ], is not supported by the production data in the present study.

The perception data in the present study also show no confusion among the three place categories of Mandarin sibilants for Cantonese speakers, although the three sets of Mandarin sibilants correspond to one set of sibilants in Cantonese. All Cantonese speakers performed very well in the perception experiment, receiving an overall correct identification score over 90%. Among the three categories of Mandarin sibilants, the correct identification rate is relatively higher for the retroflex [ʈʂ, ʈʂʰ, ʂ] - the 'new' phones, followed by the alveolo-palatal [ʨ, ʨʰ, ɕ] - the 'new similar' phones and the denti-alveolar [ʦ, ʦʰ, s] - the 'similar' phones for Cantonese speakers. This perception order of the Mandarin sibilants is in agreement with the prediction of the

Speech Learning Model proposed by Flege (1995). However, the descending order of the overall correct rate of [tʂ, tʂʰ, ʂ] > [tɕ, tɕʰ, ɕ] > [ts, tsʰ, s] is not observed in the production experiment of the present study, indicating that the L2 acquisition of Mandarin sibilants for Cantonese speakers is not wholly predicted the Speech Learning Model.

In the present study, Cantonese speakers' overall performance in perception is better than in production, indicating a difference in competence between production and perception of L2 sounds. In both production and perception, the performance of Cantonese speakers is relatively weaker in the denti-alveolar [ts, tsʰ, s] than the retroflex [tʂ, tʂʰ, ʂ] and alveolo-palatal [tɕ, tɕʰ, ɕ]. This demonstrates a similarity in L2 acquisition between production and perception.

In conclusion, the present study has presented the data on both the production and perception of Mandarin sibilants in L2. Hopefully, the findings can pave the way for further investigation of the acquisition of the Mandarin sounds in other L2 learners.

6. References

Behrens, S. J., and Blumstein, S. E. (1988). Acoustic characteristics of English voiceless fricatives. Journal of Phonetics, 16, 295-298.

Cheng, Ting Au (鄭定歐) (1995).〈二論面向九七的香港推普工作〉,《一九九七與香港中國語文研討會論文集》. 香港: 香港中國語文學會, 419-422.

Chung, Raung Fu (鍾榮富) (2009). 《對比分析與華語教學》. 台北: 正中書局, 142-154.

Chung, Raung Fu and Si, Qiu Xue (鍾榮富, 司秋雪) (2009).〈從聲學分析探討初韓日越學習者的華語擦音〉,《第九屆世界華語文教學研討會論文集》, 第一冊語言分析(1). 台北: 世界華語文教育學會, 104-123.

Evers, Vincent, Reetz, Henning, and Lahiri, Aditi. (1998). Crosslinguistic acoustic categorization of sibilants independent of phonological status. Journal of Phonetics, 26, 345-370.

Flege, James E. (1995). Second language speech learning: Theories, findings, and problems. In W. Strange (Ed.), Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Timonium: York Press, 233-272.

Hein, J. M. and Stevens, K. M. (1961). On the properties of voiceless fricative consonants. Journal of the Acoustical Society of America, 33, 589-596.

Hon, Wai Sun (韓維新) (2003).〈關於香港學生學習普通話發音難點的研究〉,《第七屆世界華語文教學研討會論文集》, 第六冊教學應用組. 台北: 世界華語文教育學會, 63-80.

Ng, Lai Ping (吳麗萍) (2001).〈如何在普通話語音學習上精益求精〉,《中國語文及人文素質教育》. 香港: 香港教育學院, 185-197.

Nowak, P. M. (2006). The role of vowel transitions and frication noise in the perception of Polish sibilants. Journal of Phonetics, 34, 139-152.

Ladefoged, P. and Johnson, K. (2015). A Course in Phonetics (7th ed.). U.S.A.: Cengage Learning, 17.

Lai, Y. H. (2009). Asymmetry in Mandarin affricate perception by learners of Mandarin Chinese, Language and Cognitive Processes, 24 (7-8), 1265-1285. DOI: 10.1080/01690960802113850.

Li, Fei (李斐) (2009).〈香港大學生普通話口語會話語音偏誤分析〉,《首都師範大學學報》, S3, 80-86.

Lee, C.Y., Zhang, Y. and Li, X.M. (2014). Acoustic characteristics of voiceless fricatives in Mandarin Chinese. Journal of Chinese Linguistics, 42 (1), 150-171.

Lee, S. I. (2011). Spectral analysis of Mandarin Chinese sibilant fricatives. In the Proceedings of the 17th International Congress of Phonetic Sciences. Hong Kong: City University of Hong Kong, 1178-1181.

Li, S.P. and Gu, W.T. (2015). Acoustic analysis of Mandarin affricates. Paper presented at the Interspeech 2015, Germany. Retrieved from http://www.researchgate.net/publication/280802627_Acoustic_Analysis_of_Mandarin_Affricates

Lee-Wong Ping (李黃萍) (2013).〈語音偏誤分析—以香港學生為例〉,《深化語文教育改革的思考與實踐》. 南京: 東南大學出版社, 277-286.

Lee, W. S. and Zee, E. (2003). Standard Chinese (Beijing). Journal of the International Phonetic Association, 33(1), 109-112.

Svantesson, J.O. (1986). Acoustic analysis of Chinese fricatives and affricates. Journal of Chinese Linguistics 14 (1), 52-70.

Wong, C. K. (2015). An acoustic analysis of Mandarin sibilants produced by Cantonese Speakers. BA thesis, City University of Hong Kong. Retrieved from CityU Institutional Repository.

Wu, Chen Huei and Su, Jui Ya (吳貞慧, 蘇睿亞) (2014).〈澳門華語學習者華語語音偏誤類型之研究〉,《國立臺北教育大學語文集刊》, 26, 103-147.

Zee, E. (1999). Chinese (Hong Kong Cantonese). Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet. Cambridge: Cambridge University Press, 58-60.

Zhu, Xin Hua, Chan, Shui Duen, and Wen, Hong Bo (祝新華, 陳瑞端, 溫紅博) (2012).〈十年間香港大學生普通話水平的發展變化——基於香港理工大學的十年考試數據〉,《語言文字應用》, 1, 58-66.

Appendix 1: Frequency values of the noise range and noise peak for the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ] for a Mandarin speaker and eight Cantonese speakers.

(a) Mandarin Speaker

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	8,181	12,169	10,999	**[ʦʰ]**	1	8,107	11,505	8,279	**[s]**	1	7,886	10,323	8,911
	2	8,107	12,095	10,292		2	8,033	11,948	11,694		2	8,107	10,618	9,000
	3	8,033	14,680	10,551		3	7,516	12,686	10,323		3	8,255	12,908	11,405
	Mean	**8,107**	**12,981**	**10,614**		Mean	**7,885**	**12,046**	**10,099**		Mean	**8,083**	**11,283**	**9,772**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʈʂ]	1	1,903	10,175	3,607	**[ʈʂʰ]**	1	2,347	9,584	3,543	**[ʂ]**	1	1,830	10,249	4,182
	2	1,903	9,954	4,460		2	2,420	9,363	3,723		2	1,830	9,289	3,226
	3	2,199	10,249	2,758		3	2,199	10,249	3,613		3	1,903	10,840	3,725
	Mean	**2,002**	**10,126**	**3,608**		Mean	**2,322**	**9,732**	**3,626**		Mean	**1,854**	**10,126**	**3,711**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	6,093	10,914	6,328	**[ʨʰ]**	1	6,113	9,363	7,991	**[ɕ]**	1	6,039	9,732	7,844
	2	6,482	10,101	7,125		2	5,448	9,732	6,821		2	6,482	10,766	6,927
	3	5,670	11,283	6,603		3	5,522	9,954	6,939		3	5,818	9,584	6,127
	Mean	**6,082**	**10,766**	**6,685**		Mean	**5,694**	**9,683**	**7,250**		Mean	**6,113**	**10,027**	**6,966**

(b) Cantonese Male 1

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	7,780	11,385	8,290	**[ʦʰ]**	1	7,689	10,439	9,949	**[s]**	1	7,419	11,881	9,554
	2	7,644	11,250	10,937		2	7,509	10,348	9,277		2	7,509	9,582	9,428
	3	7,915	11,295	8,580		3	3,769	11,115	5,324		3	7,870	10,529	9,689
	Mean	**7,780**	**11,310**	**9,269**		Mean	**6,322**	**10,634**	**8,183**		Mean	**7,599**	**10,664**	**9,557**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʧ]	1	4,399	10,303	8,878	**[ʧʰ]**	1	3,183	9,943	5,761	**[ʃ]**	1	2,281	11,070	6,797
	2	2,507	11,881	4,097		2	7,735	11,070	9,741		2	2,236	11,205	9,326
	3	2,281	11,115	6,294		3	7,013	9,582	9,469		3	2,416	11,205	5,480
	Mean	**3,062**	**11,100**	**6,423**		Mean	**5,977**	**10,198**	**8,324**		Mean	**2,311**	**11,160**	**7,201**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	7,239	11,701	8,867	**[ʨʰ]**	1	6,563	11,746	9,465	**[ɕ]**	1	7,329	10,439	8,938
	2	7,239	11,340	8,120		2	6,788	9,312	8,227		2	7,149	9,672	9,093
	3	7,013	10,889	9,546		3	7,509	11,655	10,346		3	6,923	11,430	8,657
	Mean	**7,164**	**11,310**	**8,844**		Mean	**6,953**	**10,904**	**9,346**		Mean	**7,134**	**10,514**	**8,896**

(c) Cantonese Male 2

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	7,064	12,984	8,289	**[ʦʰ]**	1	7,898	10,983	8,793	**[s]**	1	7,773	11,983	10,108
	2	7,398	12.192	7,773		2	7,356	11,108	7,949		2	7,231	11,108	8,512
	3	7,481	11,358	9,610		3	7,439	11,525	8,937		3	7,439	12,900	8,290
	Mean	**7,314**	**8,118**	**8,557**		Mean	**7,564**	**11,205**	**8,560**		Mean	**7,481**	**11,997**	**8,970**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʧ]	1	6,689	12,317	8,782	**[ʧʰ]**	1	7,314	11,900	8,824	**[ʃ]**	1	7,481	11,608	8,980
	2	8,273	10,649	9,064		2	7,148	10,441	8,280		2	7,648	12,734	8,338
	3	8,190	12,650	8,958		3	7,148	11,108	8,332		3	7,356	11,983	8,288
	Mean	**7,717**	**11,872**	**8,935**		Mean	**7,203**	**11,150**	**8,479**		Mean	**7,495**	**12,108**	**8,535**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	7,106	10,941	9,739	**[ʨʰ]**	1	7,022	11,733	8,239	**[ɕ]**	1	6,606	12,275	8,225
	2	7,356	11,608	7,678		2	6,939	12,359	8,120		2	7,481	11,733	8,156
	3	7,398	10,149	8,340		3	7,398	12,734	8,193		3	7,231	11,275	8,374
	Mean	**7,287**	**10,899**	**8,586**		Mean	**7,120**	**12,275**	**8,184**		Mean	**7,106**	**11,761**	**8,252**

(d) Cantonese Male 3

Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	7,106	11,692	9,451	**[ʦʰ]**	1	7,314	15,485	8,145	**[s]**	1	2,145	15,068	5,371
	2	6,897	13,609	8,486		2	7,398	12,400	9,326		2	2,145	12,942	9,059
	3	7,106	13,359	9,756		3	7,481	11,483	8,289		3	3,145	13,526	9,461
	Mean	**7,036**	**12,887**	**9,231**		Mean	**7,398**	**13,123**	**8,587**		Mean	**2,478**	**13,845**	**7,964**

Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tʂ]	1	2,228	14,526	9,044	**[tʂʰ]**	1	6,897	13,526	9,311	**[ʂ]**	1	6,926	15,661	8,472
	2	2,020	13,151	3,256		2	7,196	11,108	9,390		2	7,630	12,533	8,308
	3	6,772	13,442	9,778		3	2,478	14,568	9,625		3	7,379	14,859	9,567
	Mean	**3,673**	**13,706**	**7,359**		Mean	**5,524**	**13,067**	**9,442**		Mean	**7,312**	**14,351**	**8,782**

Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tɕ]	1	7,441	15,991	10,805	**[tɕʰ]**	1	7,083	14,140	8,247	**[ɕ]**	1	6,956	12,550	8,482
	2	7,737	14,926	9,332		2	7,401	13,504	8,499		2	7,210	12,614	8,113
	3	7,274	13,313	8,451		3	7,274	13,504	9,904		3	7,401	11,025	8,688
	Mean	**7,484**	**14,743**	**9,529**		Mean	**7,253**	**13,716**	**8,883**		Mean	**7,189**	**12,063**	**8,428**

(e) Cantonese Male 4

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	7,337	13,059	9,797	**[ʦʰ]**	1	7,464	12,868	11,780	**[s]**	1	7,591	12,550	9,650
	2	7,337	10,961	9,922		2	7,337	12,169	10,129		2	7,909	10,770	9,361
	3	7,464	11,088	8,966		3	7,464	11,470	9,075		3	7,591	10,579	9,511
	Mean	**7,379**	**11,703**	**9,562**		Mean	**7,422**	**12,169**	**10,328**		Mean	**7,697**	**11,300**	**9,507**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʧ]	1	2,315	10,452	7,293	**[ʧʰ]**	1	2,378	10,643	9,060	**[ʃ]**	1	1,933	10,262	3,375
	2	2,315	12,042	3,556		2	2,124	10,834	3,032		2	2,950	10,134	4,184
	3	2,505	10,389	6,215		3	2,187	9,944	6,778		3	1,870	11,597	4,345
	Mean	**2,378**	**10,961**	**5,688**		Mean	**2,230**	**10,474**	**6,290**		Mean	**2,251**	**10,664**	**3,968**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	7,687	10,957	9,893	**[ʨʰ]**	1	6,511	11,533	8,164	**[ɕ]**	1	6,956	12,105	9,864
	2	7,463	12,167	9,409		2	6,638	10,897	7,635		2	7,146	10,961	9,793
	3	7,464	12,169	10,485		3	7,337	12,805	8,706		3	7,083	11,470	10,441
	Mean	**7,538**	**11,764**	**9,929**		Mean	**6,829**	**11,745**	**8,168**		Mean	**7,062**	**11,512**	**10,033**

(f) Cantonese Female 1

Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,478	10,689	9,089	**[ʦʰ]**	1	4,910	10,599	9,131	**[s]**	1	3,924	9,569	4,590
	2	6,523	10,106	8,912		2	6,657	10,330	9,263		2	4,193	11,719	5,828
	3	6,478	10,330	9,580		3	5,716	10,957	6,686		3	4,193	10,420	6,231
	Mean	**6,493**	**10,375**	**9,194**		Mean	**5,761**	**10,629**	**8,360**		Mean	**4,103**	**10,569**	**5,550**

Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tʂ]	1	6,657	9,972	9,591	**[tʂʰ]**	1	3,208	10,689	3,695	**[ʂ]**	1	2,625	7,329	3,614
	2	6,702	10,151	9,658		2	6,478	11,047	9,793		2	4,238	10,196	6,345
	3	5,492	10,554	6,453		3	3,028	8,359	3,385		3	3,656	10,017	5,913
	Mean	**6,284**	**10,226**	**8,567**		Mean	**4,238**	**10,032**	**5,624**		Mean	**3,506**	**9,181**	**5,291**

Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak	Target sibilant	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tɕ]	1	6,030	9,434	8,918	**[tɕʰ]**	1	6,343	11,405	9,069	**[ɕ]**	1	5,537	10,151	6,960
	2	6,657	9,569	9,121		2	6,343	9,972	8,725		2	5,268	10,554	5,470
	3	6,612	10,330	6,967		3	6,567	11,719	9,779		3	5,268	9,658	5,358
	Mean	**6,433**	**9,778**	**8,335**		Mean	**6,418**	**11,032**	**9,191**		Mean	**5,358**	**10,121**	**5,929**

(g) Cantonese Female 2

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,230	12,067	9,320	**[ʦʰ]**	1	3,437	8,315	4,023	**[s]**	1	6,731	11,942	9,586
	2	6,814	11,233	9,141		2	5,939	11,858	8,926		2	6,612	9,927	9,419
	3	6,355	11,608	10,071		3	6,272	9,732	6,566		3	6,612	11,674	9,667
	Mean	**6,466**	**11,636**	**9,511**		Mean	**5,216**	**9,968**	**6,505**		Mean	**6,652**	**11,181**	**9,557**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʧ]	1	6,970	10,778	9,631	**[ʧʰ]**	1	6,657	10,733	8,423	**[ʃ]**	1	6,791	9,613	8,556
	2	7,060	10,106	7,146		2	6,523	11,360	6,978		2	6,254	11,719	9,452
	3	7,374	12,436	11,509		3	6,970	10,957	9,151		3	7,508	11,092	9,495
	Mean	**7,135**	**11,107**	**9,429**		Mean	**6,717**	**11,017**	**8,184**		Mean	**6,851**	**10,808**	**9,168**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	6,209	10,017	9,300	**[ʨʰ]**	1	7,150	11,584	9,650	**[ɕ]**	1	6,657	12,077	9,022
	2	6,343	10,823	9,455		2	6,702	11,226	9,989		2	6,343	10,644	9,655
	3	6,209	9,748	9,248		3	6,254	11,181	7,652		3	6,254	9,613	9,754
	Mean	**6,254**	**10,196**	**9,334**		Mean	**6,702**	**11,330**	**9,097**		Mean	**6,418**	**10,778**	**9,477**

(h) Cantonese Female 3

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	8,225	12,884	9,305	**[ʦʰ]**	1	7,284	10,913	9,842	**[s]**	1	7,329	11,764	9,572
	2	7,822	12,346	8,343		2	7,239	12,973	8,654		2	7,463	11,808	9,516
	3	7,553	10,330	7,976		3	7,239	10,017	9,744		3	7,956	12,212	9,668
	Mean	**7,867**	**11,853**	**8,541**		Mean	**7,254**	**11,301**	**9,413**		Mean	**7,583**	**11,928**	**9,585**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ̢]	1	7,777	11,629	8,008	**[ʦ̢ʰ]**	1	6,926	9,748	6,974	**[ʂ]**	1	6,030	10,868	6,597
	2	7,598	11,226	9,903		2	7,239	10,465	9,426		2	6,343	11,405	6,412
	3	7,015	11,047	9,278		3	7,866	11,674	10,537		3	6,164	11,764	6,443
	Mean	**7,463**	**11,301**	**9,063**		Mean	**7,344**	**10,629**	**8,979**		Mean	**6,179**	**11,346**	**6,484**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	6,164	11,540	9,623	**[ʨʰ]**	1	6,791	11,988	9,526	**[ɕ]**	1	7,239	12,839	9,926
	2	6,881	11,540	7,636		2	7,418	11,584	9,229		2	6,523	12,077	9,236
	3	7,015	11,584	8,580		3	6,836	11,853	9,749		3	6,523	10,778	7,819
	Mean	**6,687**	**11,555**	**8,613**		Mean	**7,015**	**11,808**	**9,501**		Mean	**6,762**	**11,898**	**8,994**

(i) Cantonese Female 4

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	7,822	12,525	9,190	**[ʦʰ]**	1	7,194	10,196	9,738	**[s]**	1	9,492	12,647	9,632
	2	7,239	11,629	9,287		2	7,015	10,689	9,582		2	8,771	12,016	9,029
	3	7,553	10,599	9,298		3	6,878	11,881	9,186		3	8,456	12,512	9,509
	Mean	**7,538**	**11,584**	**9,258**		Mean	**7,029**	**10,922**	**9,502**		Mean	**8,906**	**12,392**	**9,390**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tʂ]	1	3,453	9,312	3,481	**[tʂʰ]**	1	2,777	10,213	3,388	**[ʂ]**	1	2,507	9,402	3,272
	2	2,552	9,763	8,746		2	2,777	9,763	6,648		2	2,597	9,763	4,731
	3	2,507	9,177	3,507		3	3,363	9,672	3,559		3	2,507	9,402	3,003
	Mean	**2,837**	**9,417**	**5,245**		Mean	**2,972**	**9,883**	**4,532**		Mean	**2,537**	**9,522**	**3,669**

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tɕ]	1	4,715	10,844	8,184	**[tɕʰ]**	1	5,526	9,222	6,292	**[ɕ]**	1	5,301	10,484	5,368
	2	4,940	10,078	6,511		2	5,706	10,033	6,180		2	5,346	9,627	9,250
	3	5,211	10,394	5,962		3	5,211	8,816	5,960		3	6,788	10,484	7,082
	Mean	**4,955**	**10,439**	**6,886**		Mean	**5,481**	**9,357**	**6,144**		Mean	**5,812**	**10,198**	**7,233**

(j) Four Cantonese male speakers (mean of 12 tokens = 3 tokens x 4 speakers)

Target sibilant	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	7,780	11,310	9,269	**[ʦʰ]**	1	6,322	10,634	8,183	**[s]**	1	7,599	10,664	9,557
	2	7,314	8,118	8,557		2	7,564	11,205	8,560		2	7,481	11,997	8,970
	3	7,036	12,887	9,231		3	7,398	13,123	8,587		3	2,478	13,845	7,964
	4	7,379	11,703	9,562		4	7,422	12,169	10,328		4	7,697	11,300	9,507
	Mean	**7,377**	**11,005**	**9,155**		Mean	**7,177**	**11,783**	**8,915**		Mean	**6,314**	**11,952**	**9,000**

Target sibilant	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ̢]	1	3,062	11,100	6,423	**[ʦ̢ʰ]**	1	5,977	10,198	8,324	**[ʂ]**	1	2,311	11,160	7,201
	2	7,717	11,872	8,935		2	7,203	11,150	8,479		2	7,495	12,108	8,535
	3	3,673	13,706	7,359		3	5,524	13,067	9,442		3	7,312	14,351	8,782
	4	2,378	10,961	5,688		4	2,230	10,474	6,290		4	2,251	10,664	3,968
	Mean	**4,208**	**11,910**	**7,101**		Mean	**5,234**	**11,222**	**8,134**		Mean	**4,842**	**12,071**	**7,122**

Target sibilant	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	7,164	11,310	8,844	**[ʨʰ]**	1	6,953	10,904	9,346	**[ɕ]**	1	7,134	10,514	8,896
	2	7,287	10,889	8,586		2	7,120	12,275	8,184		2	7,106	11,761	8,252
	3	7,484	14,743	9,529		3	7,253	13,716	8,883		3	7,189	12,063	8,428
	4	7,538	11,764	9,929		4	6,829	11,745	8,168		4	7,062	11,512	10,033
	Mean	**7,368**	**12,177**	**9,222**		Mean	**7,039**	**12,160**	**8,645**		Mean	**7,123**	**11,463**	**8,902**

(k) Four Cantonese female speakers (mean of 12 tokens = 3 tokens x 4 speakers)

Target sibilant	Female no.	Noise range		Noise peak	**Target sibilant**	Female no.	Noise range		Noise peak	**Target sibilant**	Female no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,493	10,375	9,194	**[ʦʰ]**	1	5,761	10,629	8,360	**[s]**	1	4,103	10,569	5,550
	2	6,466	11,636	9,511		2	5,216	9,968	6,505		2	6,652	11,181	9,557
	3	7,867	11,853	8,541		3	7,254	11,301	9,413		3	7,583	11,928	9,585
	4	7,538	11,584	9,258		4	7,029	10,922	9,502		4	8,906	12,392	9,390
	Mean	**7,091**	**11,362**	**9,126**		Mean	**6,315**	**10,705**	**8,445**		Mean	**6,811**	**11,518**	**8,521**

Target sibilant	Female no.	Noise range		Noise peak	**Target sibilant**	Female no.	Noise range		Noise peak	**Target sibilant**	Female no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[tʂ]	1	6,284	10,226	8,567	**[tʂʰ]**	1	4,238	10,032	5,624	**[ʂ]**	1	3,506	9,181	5,291
	2	7,135	11,107	9,429		2	6,717	11,017	8,184		2	6,851	10,808	9,168
	3	7,463	11,301	9,063		3	7,344	10,629	8,979		3	6,179	11,346	6,484
	4	2,837	9,417	5,245		4	2,972	9,883	4,532		4	2,537	9,522	3,669
	Mean	**5,930**	**10,513**	**8,076**		Mean	**5,318**	**10,390**	**6,830**		Mean	**4,768**	**10,214**	**6,153**

Target sibilant	Female no.	Noise range		Noise peak	**Target sibilant**	Female no.	Noise range		Noise peak	**Target sibilant**	Female no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʨ]	1	6,433	9,778	8,335	**[ʨʰ]**	1	6,418	11,032	9,191	**[ɕ]**	1	5,358	10,121	5,929
	2	6,254	10,196	9,334		2	6,702	11,330	9,097		2	6,418	10,778	9,477
	3	6,687	11,555	8,613		3	7,015	11,808	9,501		3	6,762	11,898	8,994
	4	4,955	10,439	6,886		4	5,481	9,357	6,144		4	5,812	10,198	7,233
	Mean	**6,082**	**10,492**	**8,292**		Mean	**6,404**	**10,882**	**8,483**		Mean	**6,088**	**10,749**	**7,908**

Appendix 2: Percentages of correct and incorrect production of the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ] for eight Cantonese speakers.

(a) Cantonese Male 1 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [ʈʂ, ʈʂʰ, ʂ]	Alveolo-palatal [ʨ ʨʰ ɕ]	New form
[ʦ]	2/3	0	0	1/3
[ʦʰ]	2/3	0	0	1/3
[s]	3/3	0	0	0
Overall	**7/9 (77.78%)**	**0**	**0**	**2/9 (22.22%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ʦ, ʦʰ, s]	Alveolo-palatal [ʨ ʨʰ ɕ]	New form
[ʈʂ]	1/3	0	0	2/3
[ʈʂʰ]	0	1/3	0	2/3
[ʂ]	0	0	0	3/3
Overall	**1/9 (11.11%)**	**1/9 (11.11%)**	**0**	**7/9 (77.78%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ʦ, ʦʰ, s]	Retroflex [ʈʂ, ʈʂʰ, ʂ]	New form
[ʨ]	0	0	0	3/3
[ʨʰ]	0	1/3	0	2/3
[ɕ]	0	1/3	0	2/3
Overall	**0**	**2/9 (22.22%)**	**0**	**7/9 (77.78%)**

(b) Cantonese Male 2 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [ʈʂ, ʈʂʰ, ʂ]	Alveolo-palatal [ʨ ʨʰ ɕ]	New form
[ʦ]	2/3	0	0	1/3
[ʦʰ]	2/3	0	0	1/3
[s]	2/3	0	0	1/3
Overall	**6/9 (66.67%)**	**0**	**0**	**3/9 (33.33%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ʦ, ʦʰ, s]	Alveolo-palatal [ʨ ʨʰ ɕ]	New form
[ʈʂ]	0	2/3	0	1/3
[ʈʂʰ]	0	1/3	0	2/3
[ʂ]	0	2/3	0	1/3
Overall	**0**	**5/9 (55.56%)**	**0**	**4/9 (44.44%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ʦ, ʦʰ, s]	Retroflex [ʈʂ, ʈʂʰ, ʂ]	New form
[ʨ]	0	2/3	0	1/3
[ʨʰ]	0	0	0	3/3
[ɕ]	0	1/3	0	2/3
Overall	**0**	**3/9 (33.33%)**	**0**	**6/9 (66.67%)**

(c) Cantonese Male 3 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[ts]	0	0	0	3/3
[tsʰ]	1/3	0	0	2/3
[s]	0	0	0	3/3
Overall	**1/9 (11.11%)**	**0**	**0**	**8/9 (88.89%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[tʂ]	1/3	0	0	2/3
[tʂʰ]	0	0	0	3/3
[ʂ]	0	2/3	0	1/3
Overall	**1/9 (11.11%)**	**2/9 (22.22%)**	**0**	**6/9 (66.67%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Retroflex [tʂ, tʂʰ, ʂ]**	**New form**
[tɕ]	0	2/3	0	1/3
[tɕʰ]	0	1/3	0	2/3
[ɕ]	0	1/3	0	2/3
Overall	**0**	**4/9 (44.44%)**	**0**	**5/9 (55.56%)**

(d) Cantonese Male 4 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[ts]	3/3	0	0	0
[tsʰ]	3/3	0	0	0
[s]	3/3	0	0	0
Overall	**9/9 (100%)**	**0**	**0**	**0**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[tʂ]	1/3	0	0	2/3
[tʂʰ]	1/3	0	0	2/3
[ʂ]	3/3	0	0	0
Overall	**5/9 (55.56%)**	**0**	**0**	**4/9 (44.44%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Retroflex [tʂ, tʂʰ, ʂ]**	**New form**
[tɕ]	0	3/3	0	0
[tɕʰ]	1/3	1/3	0	1/3
[ɕ]	0	0	0	3/3
Overall	**1/9 (11.11%)**	**4/9 (44.44%)**	**0**	**4/9 (44.44%)**

(e) Cantonese Female 1 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	Alveolo-palatal [tɕ tɕʰ ɕ]	New form
[ts]	0	0	0	3/3
[tsʰ]	0	0	1/3	2/3
[s]	0	1/3	0	2/3
Overall	**0**	**1/9 (11.11%)**	**1/9 (11.11%)**	**7/9 (77.78%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	Alveolo-palatal [tɕ tɕʰ ɕ]	New form
[tʂ]	0	0	1/3	2/3
[tʂʰ]	2/3	0	0	1/3
[ʂ]	1/3	0	0	2/3
Overall	**3/9 (33.33%)**	**0**	**1/9 (11.11%)**	**5/9 (55.56%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	Retroflex [tʂ, tʂʰ, ʂ]	New form
[tɕ]	1/3	0	0	2/3
[tɕʰ]	0	0	0	3/3
[ɕ]	3/3	0	0	0
Overall	**4/9 (44.44%)**	**0**	**0**	**5/9 (55.56%)**

(f) Cantonese Female 2 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	Alveolo-palatal [tɕ tɕʰ ɕ]	New form
[ts]	0	0	0	3/3
[tsʰ]	0	1/3	1/3	1/3
[s]	0	0	0	3/3
Overall	**0**	**1/9 (11.11%)**	**1/9 (11.11%)**	**7/9 (77.78%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	Alveolo-palatal [tɕ tɕʰ ɕ]	New form
[tʂ]	0	1/3	0	2/3
[tʂʰ]	0	0	1/3	2/3
[ʂ]	0	1/3	0	2/3
Overall	**0**	**2/9 (22.22%)**	**1/9 (11.11%)**	**6/9 (66.67%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	Retroflex [tʂ, tʂʰ, ʂ]	New form
[tɕ]	0	0	0	3/3
[tɕʰ]	1/3	0	0	2/3
[ɕ]	0	0	0	3/3
Overall	**1/9 (11.11%)**	**0**	**0**	**8/9 (88.89%)**

(g) Cantonese Female 3 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[ts]	2/3	0	0	1/3
[tsʰ]	0	0	0	3/3
[s]	3/3	0	0	0
Overall	**5/9 (55.56%)**	**0**	**0**	**4/9 (44.44%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[tʂ]	0	2/3	0	1/3
[tʂʰ]	0	1/3	0	2/3
[ʂ]	0	0	3/3	0
Overall	**0**	**3/9 (33.33%)**	**3/9 (33.33%)**	**3/9 (33.33%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Retroflex [tʂ, tʂʰ, ʂ]**	**New form**
[tɕ]	1/3	0	0	2/3
[tɕʰ]	0	1/3	0	2/3
[ɕ]	0	0	0	3/3
Overall	**1/9 (11.11%)**	**1/9 (11.11%)**	**0**	**7/9 (77.78%)**

(h) Cantonese Female 4 (3 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[ts]	2/3	0	0	1/3
[tsʰ]	0	0	0	3/3
[s]	3/3	0	0	0
Overall	**5/9 (55.56%)**	**0**	**0**	**4/9 (44.44%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Alveolo-palatal [tɕ tɕʰ ɕ]**	**New form**
[tʂ]	2/3	0	0	1/3
[tʂʰ]	2/3	0	0	1/3
[ʂ]	3/3	0	0	0
Overall	**7/9 (77.78%)**	**0**	**0**	**2/9 (22.22%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	**Retroflex [tʂ, tʂʰ, ʂ]**	**New form**
[tɕ]	2/3	0	0	1/3
[tɕʰ]	3/3	0	0	0
[ɕ]	2/3	0	0	1/3
Overall	**7/9 (77.78%)**	**0**	**0**	**2/9 (22.22%)**

(i) Eight Cantonese speakers (24 tokens for each sibilant)

Target sibilant	Correct	Incorrect		
		Retroflex [tʂ, tʂʰ, ʂ]	Alveolo-palatal [tɕ tɕʰ ɕ]	New form
[ts]	11/24	0	0	13/24
[tsʰ]	8/24	1/24	2/24	13/24
[s]	14/24	1/24	0	9/24
Overall	**33/72 (45.83%)**	**2/72 (2.78%)**	**2/72 (2.78%)**	**35/72 (48.61%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	Alveolo-palatal [tɕ tɕʰ ɕ]	New form
[tʂ]	5/24	5/24	1/24	13/24
[tʂʰ]	5/24	3/24	1/24	15/24
[ʂ]	7/24	5/24	3/24	9/24
Overall	**17/72 (23.61%)**	**13/72 (18.06%)**	**5/72 (6.94%)**	**37/72 (51.39%)**

Target sibilant	Correct	Incorrect		
		Denti-alveolar [ts, tsʰ, s]	Retroflex [tʂ, tʂʰ, ʂ]	New form
[tɕ]	4/24	7/24	0	13/24
[tɕʰ]	5/24	4/24	0	15/24
[ɕ]	5/24	3/24	0	16/24
Overall	**14/72 (19.44%)**	**14/72 (19.44%)**	**0**	**44/72 (61.11%)**

Appendix 3: Frequency values of the noise range and noise peak for the Cantonese sibilants [ts, tsʰ, s] for eight Cantonese speakers.

(a) Cantonese Male 1

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ts]	1	6,209	10,017	8,764	**[tsʰ]**	1	6,075	11,450	9,120	**[s]**	1	6,299	11,405	6,762
	2	6,612	11,181	8,865		2	6,612	10,913	9,795		2	7,060	11,316	8,440
	3	6,836	11,360	8,903		3	6,791	11,405	9,703		3	6,970	11,764	8,482
	Mean	**6,552**	**10,853**	**8,703**		Mean	**6,493**	**11,256**	**9,539**		Mean	**6,776**	**11,495**	**7,895**

(b) Cantonese Male 2

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ts]	1	6,657	13,242	8,896	**[tsʰ]**	1	6,791	11,360	8,489	**[s]**	1	6,612	12,570	8,292
	2	6,836	10,778	8,265		2	6,747	12,525	8,574		2	6,702	12,660	9,161
	3	6,523	12,570	9,693		3	6,791	11,360	8,904		3	6,836	12,391	9,514
	Mean	**6,672**	**12,197**	**8,951**		Mean	**6,776**	**11,748**	**8,656**		Mean	**6,717**	**12,540**	**8,989**

(c) Cantonese Male 3

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ts]	1	6,299	13,824	7,671	**[tsʰ]**	1	7,642	11,988	9,046	**[s]**	1	6,254	10,689	8,966
	2	6,657	12,346	8,648		2	6,299	11,316	7,902		2	6,164	12,749	8,353
	3	6,523	10,509	9,600		3	6,523	13,555	8,805		3	6,433	12,167	7,091
	Mean	**6,493**	**12,226**	**8,640**		Mean	**6,821**	**12,286**	**8,584**		Mean	**6,284**	**11,868**	**8,137**

(c) Cantonese Male 4

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,926	11,047	9,520	**[ʦʰ]**	1	6,747	12,301	7,106	**[s]**	1	7,194	11,002	9,687
	2	7,418	10,957	10,201		2	6,881	12,749	8,808		2	6,926	13,108	7,588
	3	7,015	11,092	8,256		3	7,418	12,167	8,606		3	6,657	12,660	9,143
	Mean	**7,120**	**11,032**	**9,326**		Mean	**7,015**	**12,406**	**8,173**		Mean	**6,926**	**12,257**	**8,806**

(e) Cantonese Female 1

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,343	14,093	9,210	**[ʦʰ]**	1	7,015	13,466	8,645	**[s]**	1	6,119	13,063	8,346
	2	6,164	13,600	7,976		2	6,388	13,287	9,470		2	6,791	13,511	9,525
	3	6,791	11,405	9,534		3	6,791	11,360	8,063		3	6,567	13,421	8,709
	Mean	**6,433**	**13,033**	**8,907**		Mean	**6,731**	**12,704**	**8,726**		Mean	**6,492**	**13,332**	**8,860**

(f) Cantonese Female 2

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	5,985	9,927	9,574	**[ʦʰ]**	1	6,299	10,868	10,061	**[s]**	1	6,704	11,357	9,505
	2	6,523	11,540	8,449		2	6,433	11,629	8,986		2	6,482	11,357	8,460
	3	6,523	12,301	9,121		3	6,567	11,450	9,434		3	6,209	12,794	9,549
	Mean	**6,344**	**11,256**	**9,048**		Mean	**6,433**	**11,316**	**9,494**		Mean	**6,465**	**11,836**	**9,171**

(g) Cantonese Female 3

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,836	14,631	9,654	**[ʦʰ]**	1	6,881	14,317	9,495	**[s]**	1	6,657	13,914	8,339
	2	7,194	14,631	8,844		2	6,030	12,256	9,375		2	6,747	11,002	9,580
	3	6,612	13,376	9,703		3	6,433	12,167	8,878		3	6,567	11,271	7,993
	Mean	**6,881**	**14,213**	**10,547**		Mean	**6,448**	**12,913**	**9,249**		Mean	**6,657**	**12,062**	**8,637**

(h) Cantonese Female 4

Target sibilant	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak	**Target sibilant**	Token no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,881	10,913	9,511	**[ʦʰ]**	1	7,147	9,584	9,256	**[s]**	1	7,329	11,450	9,708
	2	6,791	11,540	10,741		2	6,999	9,584	8,551		2	6,209	11,360	9,531
	3	7,194	11,002	9,163		3	6,999	9,806	9,698		3	6,747	10,375	9,507
	Mean	**6,955**	**11,152**	**9,805**		Mean	**7,048**	**9,658**	**9,168**		Mean	**6,762**	**11,062**	**9,582**

(i) Four Cantonese male speakers (mean of tokens = 3 tokens x 4 speakers)

Target sibilant	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
[ʦ]	1	6,552	10,853	8,703	**[ʦʰ]**	1	6,493	11,256	9,539	**[s]**	1	6,776	11,495	7,895
	2	6,672	12,197	8,951		2	6,776	11,748	8,656		2	6,717	12,540	8,989
	3	6,493	12,226	8,640		3	6,821	12,286	8,584		3	6,284	11,868	8,137
	4	7,120	11,032	9,326		4	7,015	12,406	8,173		4	6,926	12,257	8,806
	Mean	**6,709**	**11,577**	**8,905**		Mean	**6,776**	**11,924**	**8,738**		Mean	**6,676**	**12,040**	**8,457**

(j) Four Cantonese female speakers (mean of tokens = 3 tokens x 4 speakers)

Target sibilant	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak	**Target sibilant**	Male no.	Noise range		Noise peak
		Min value	Max value				Min value	Max value				Min value	Max value	
	1	6,433	13,033	8,907		1	6,731	12,704	8,726		1	6,492	13,332	8,860
	2	6,344	11,256	9,048		2	6,433	11,316	9,494		2	6,465	11,836	9,171
[ʦ]	3	6,881	14,213	10,547	**[ʦʰ]**	3	6,448	12,913	9,249	**[s]**	3	6,657	12,062	8,637
	4	6,955	11,152	9,805		4	7,048	9,658	9,168		4	6,762	11,062	9,582
	Mean	**6,653**	**12,414**	**9,577**		Mean	**6,665**	**11,648**	**9,159**		Mean	**6,594**	**12,073**	**9,063**

Appendix 4: Perceptual assessment by Mandarin speakers for the production of Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ] from eight Cantonese speakers.

(a) Cantonese Male 1 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[ʦɿ˥]	**[ʈʂɻ˥]**	**[ʨi˥]**	**None**
[ʦɿ˥]	96.7%	3.3%	0%	0%
[ʈʂɻ˥]	0%	100%	0%	0%
[ʨi˥]	6.7%	3.3%	63.3%	26.7%

Target word	Selected word			
	[ʦʰɿ˥]	**[ʈʂʰɻ˥]**	**[ʨʰi˥]**	**None**
[ʦʰɿ˥]	26.7%	66.7%	0%	6.6%
[ʈʂʰɻ˥]	53.3%	36.7%	0%	10%
[ʨʰi˥]	0%	0%	100%	0%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	63.3%	33.3%	0%	3.3%
[ʂɻ˥]	0%	100%	0%	0%
[ɕi˥]	3.3%	0%	80%	16.7%

(b) Cantonese Male 2 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[ʦɿ˥]	**[ʈʂɻ˥]**	**[ʨi˥]**	**None**
[ʦɿ˥]	90%	6.7%	0%	3.3%
[ʈʂɻ˥]	3.3%	93.3%	0%	3.3%
[ʨi˥]	0%	6.7%	90%	3.3%

Target word	Selected word			
	[ʦʰɿ˥]	**[ʈʂʰɻ˥]**	**[ʨʰi˥]**	**None**
[ʦʰɿ˥]	96.7%	3.3%	0%	0%
[ʈʂʰɻ˥]	50%	50%	0%	0%
[ʨʰi˥]	0%	0%	96.7%	3.3%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	86.6%	6.7%	0%	6.7%
[ʂɻ˥]	10%	83.3%	0%	6.7%
[ɕi˥]	0%	0%	80%	20%

(c) Cantonese Male 3 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	6.7%	93.3%	0%	0%
[tʂɻ˥]	6.7%	93.3%	0%	0%
[tɕi˥]	0%	0%	73.3%	26.7%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɻ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	16.7%	80%	0%	3.3%
[tʂʰɻ˥]	6.7%	93.3%	0%	0%
[tɕʰi˥]	0%	0%	86.7%	13.3%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	6.7%	56.7%	0%	36.6%
[ʂɻ˥]	33.3%	66.7%	0%	0%
[ɕi˥]	0%	0%	80%	20%

(d) Cantonese Male 4 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	100%	0%	0%	0%
[tʂɻ˥]	6.7%	93.3%	0%	0%
[tɕi˥]	3.3%	0%	70%	26.7%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɻ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	100%	0%	0%	0%
[tʂʰɻ˥]	0%	100%	0%	0%
[tɕʰi˥]	0%	0%	90%	10%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	100%	0%	0%	0%
[ʂɻ˥]	0%	100%	0%	0%
[ɕi˥]	0%	0%	**63.3%**	**36.7%**

(e) Cantonese Female 1 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	93.3%	6.7%	0%	0%
[tʂɻ˥]	40%	13.3%	0%	46.7%
[tɕi˥]	0%	3.3%	93.3%	3.3%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɻ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	63.3%	33.3%	0%	3.3%
[tʂʰɻ˥]	3.3%	56.7%	33.3%	6.7%
[tɕʰi˥]	73.3%	26.7%	0%	0%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	30%	60%	0%	10%
[ʂɻ˥]	3.3%	93.3%	0%	3.3%
[ɕi˥]	0%	0%	100%	0%

(f) Cantonese Female 2 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	3.3%	96.7%	0%	0%
[tʂɻ˥]	20%	73.3%	0%	6.7%
[tɕi˥]	0%	0%	100%	0%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɻ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	6.7%	90%	0%	3.3%
[tʂʰɻ˥]	0%	96.7%	0%	3.3%
[tɕʰi˥]	0%	0%	100%	0%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	33.3%	60%	0%	6.7%
[ʂɻ˥]	16.7%	76.7%	0%	6.6%
[ɕi˥]	0%	0%	100%	0%

(g) Cantonese Female 3 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	100%	0%	0%	0%
[tʂɻ˥]	6.7%	93.3%	0%	0%
[tɕi˥]	0%	0%	93.3%	6.7%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɻ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	13.3%	16.7%	0%	70%
[tʂʰɻ˥]	43.3%	53.3%	0%	3.3%
[tɕʰi˥]	0%	0%	93.3%	6.7%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	96.7%	3.3%	0%	0%
[ʂɻ˥]	0%	93.3%	0%	6.7%
[ɕi˥]	3.3%	0%	86.7%	10%

(h) Cantonese Female 4 (30 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	100%	0%	0%	0%
[tʂɻ˥]	3.3%	96.7%	0%	0%
[tɕi˥]	0%	0%	100%	0%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɻ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	93.3%	6.7%	0%	0%
[tʂʰɻ˥]	0%	100%	0%	0%
[tɕʰi˥]	0%	0%	100%	0%

Target word	Selected word			
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	100%	0%	0%	0%
[ʂɻ˥]	0%	100%	0%	0%
[ɕi˥]	0%	0%	100%	0%

(i) Eight Cantonese speakers (240 responses for each target word; shaded area = correct production)

Target word	Selected word			
	[tsɿ˥]	**[tʂɿ˥]**	**[tɕi˥]**	**None**
[tsɿ˥]	73.7%	25.8%	0%	2.5%
[tʂɿ˥]	10.8%	82.1%%	0%	7.1%
[tɕi˥]	1.2%	1.6%	85.4%	11.6%

Target word	Selected word			
	[tsʰɿ˥]	**[tʂʰɿ˥]**	**[tɕʰi˥]**	**None**
[tsʰɿ˥]	52.1%	37.1%	0%	10.8%
[tʂʰɿ˥]	19.5%	73.3%	4.1%	2.9%
[tɕʰi˥]	9.1%	3.3%	83.3%	4.1%

Target word	Selected word			
	[sɿ˥]	**[ʂɿ˥]**	**[ɕi˥]**	**None**
[sɿ˥]	64.5%	27.5%	0%	7.9%
[ʂɿ˥]	7.9%	89.1%	0%	2.9%
[ɕi˥]	0.8%	0%	86.2%	12.9%

Appendix 5: Perception identification of the Mandarin sibilants [ʦ, ʦʰ, s], [ʈʂ, ʈʂʰ, ʂ] and [ʨ, ʨʰ, ɕ] for eight Cantonese speakers.

(a) Cantonese Male 1 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[ʦɿ˥]	**[ʈʂɻ˥]**	**[ʨi˥]**
[ʦɿ˥]	60%	40%	0%
[ʈʂɻ˥]	0%	100%	0%
[ʨi˥]	0%	0%	100%

Target word	Selected word		
	[ʦɿʰ˥]	**[ʈʂɻʰ˥]**	**[ʨʰi˥]**
[ʦɿʰ˥]	60%	40%	0%
[ʈʂɻʰ˥]	60%	40%	0%
[ʨʰi˥]	0%	0%	100%

Target word	Selected word		
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**
[sɿ˥]	100%	0%	0%
[ʂɻ˥]	0%	100%	0%
[ɕi˥]	0%	0%	100%

(b) Cantonese Male 2 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[ʦɿ˥]	**[ʈʂɻ˥]**	**[ʨi˥]**
[ʦɿ˥]	100%	0%	0%
[ʈʂɻ˥]	0%	100%	0%
[ʨi˥]	0%	0%	100%

Target word	Selected word		
	[ʦɿʰ˥]	**[ʈʂɻʰ˥]**	**[ʨʰi˥]**
[ʦɿʰ˥]	80%	20%	0%
[ʈʂɻʰ˥]	0%	100%	0%
[ʨʰi˥]	0%	0%	100%

Target word	Selected word		
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**
[sɿ˥]	100%	0%	0%
[ʂɻ˥]	0%	100%	0%
[ɕi˥]	0%	0%	100%

(c) Cantonese Male 3 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˩]	**[tʂɻ˩]**	**[tɕi˩]**
[tsɿ˩]	60%	40%	0%
[tʂɻ˩]	0%	80%	20%
[tɕi˩]	20%	0%	80%

Target word	Selected word		
	[tsɿʰ˩]	**[tʂɻʰ˩]**	**[tɕʰi˩]**
[tsɿʰ˩]	20%	60%	20%
[tʂɻʰ˩]	0%	80%	20%
[tɕʰi˩]	20%	0%	80%

Target word	Selected word		
	[sɿ˩]	**[ʂɻ˩]**	**[ɕi˩]**
[sɿ˩]	100%	0%	0%
[ʂɻ˩]	40%	60%	0%
[ɕi˩]	40%	20%	40%

(d) Cantonese Male 4 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˩]	**[tʂɻ˩]**	**[tɕi˩]**
[tsɿ˩]	100%	0%	0%
[tʂɻ˩]	0%	100%	0%
[tɕi˩]	0%	0%	100%

Target word	Selected word		
	[tsɿʰ˩]	**[tʂɻʰ˩]**	**[tɕʰi˩]**
[tsɿʰ˩]	100%	0%	0%
[tʂɻʰ˩]	0%	100%	0%
[tɕʰi˩]	0%	0%	100%

Target word	Selected word		
	[sɿ˩]	**[ʂɻ˩]**	**[ɕi˩]**
[sɿ˩]	100%	0%	0%
[ʂɻ˩]	0%	100%	0%
[ɕi˩]	0%	0%	100%

(e) Cantonese Female 1 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**
[tsɿ˥]	100%	0%	0%
[tʂɻ˥]	0%	100%	0%
[tɕi˥]	0%	20%	80%

Target word	Selected word		
	[tsɿʰ˥]	**[tʂɻʰ˥]**	**[tɕʰi˥]**
[tsɿʰ˥]	80%	0%	20%
[tʂɻʰ˥]	0%	100%	0%
[tɕʰi˥]	20%	0%	80%

Target word	Selected word		
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**
[sɿ˥]	100%	0%	0%
[ʂɻ˥]	0%	100%	0%
[ɕi˥]	0%	0%	100%

(f) Cantonese Female 2 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**
[tsɿ˥]	80%	0%	20%
[tʂɻ˥]	0%	100%	0%
[tɕi˥]	0%	0%	100%

Target word	Selected word		
	[tsɿʰ˥]	**[tʂɻʰ˥]**	**[tɕʰi˥]**
[tsɿʰ˥]	100%	0%	0%
[tʂɻʰ˥]	0%	100%	0%
[tɕʰi˥]	0%	0%	100%

Target word	Selected word		
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**
[sɿ˥]	80%	20%	0%
[ʂɻ˥]	0%	100%	0%
[ɕi˥]	0%	0%	100%

(g) Cantonese Female 3 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**
[tsɿ˥]	60%	40%	0%
[tʂɻ˥]	0%	100%	0%
[tɕi˥]	40%	0%	60%

Target word	Selected word		
	[tsɿʰ˥]	**[tʂɻʰ˥]**	**[tɕʰi˥]**
[tsɿʰ˥]	100%	0%	0%
[tʂɻʰ˥]	0%	100%	0%
[tɕʰi˥]	0%	0%	100%

Target word	Selected word		
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**
[sɿ˥]	100%	0%	0%
[ʂɻ˥]	0%	100%	0%
[ɕi˥]	0%	0%	100%

(h) Cantonese Female 4 (5 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˥]	**[tʂɻ˥]**	**[tɕi˥]**
[tsɿ˥]	100%	0%	0%
[tʂɻ˥]	0%	100%	0%
[tɕi˥]	0%	0%	100%

Target word	Selected word		
	[tsɿʰ˥]	**[tʂɻʰ˥]**	**[tɕʰi˥]**
[tsɿʰ˥]	80%	20%	0%
[tʂɻʰ˥]	0%	100%	0%
[tɕʰi˥]	0%	0%	100%

Target word	Selected word		
	[sɿ˥]	**[ʂɻ˥]**	**[ɕi˥]**
[sɿ˥]	100%	0%	0%
[ʂɻ˥]	0%	100%	0%
[ɕi˥]	0%	0%	100%

(i) Eight Cantonese speakers (40 responses for each target word; shaded area = correct identification)

Target word	Selected word		
	[tsɿ˩]	**[tʂɻ˩]**	**[tɕi˩]**
[tsɿ˩]	82.5%	15%	2.5%
[tʂɻ˩]	2.5%	97.5%	0%
[tɕi˩]	7.5%	2.5%	90%

Target word	Selected word		
	[tsɿʰ˩]	**[tʂɻʰ˩]**	**[tɕʰi˩]**
[tsɿʰ˩]	77.5%	17.5%	5%
[tʂɻʰ˩]	7.5%	90%	2.5%
[tɕʰi˩]	5%	0%	95%

Target word	Selected word		
	[sɿ˩]	**[ʂɻ˩]**	**[ɕi˩]**
[sɿ˩]	97.5%	2.5%	0%
[ʂɻ˩]	5%	95%	0%
[ɕi˩]	5%	2.5%	92.5%