香港城市大學
City University of Hong Kong

CTL 4235 Final Year Project

Topic:

Production and Perception of Cantonese Tones by Mandarin-speaking Learners

Student Name: Ku Ka Wai
Student Number:

Supervisor: Dr Lee Wai Sum
Submission Date: 20th May, 2010

CONTENT

ABSTRACT

This is a phonetic study of the production and perception of the six contrastive tones, [55, 33, 22, 21, 25, 23], in Cantonese by Mandarin-speaking learners. Three native Mandarin speakers, who have received Cantonese learning for two to three years, participated in both of the production and perception tasks of the study. In the production task, three tokens of each of the test words, [si 55] 思, [si 33] 試, [si 22] 事, [si 21] 時, [si 25] 史 and [si 23] 市, associated with the six Cantonese tones were obtained from the subjects. In the perception task, the subjects were to identify 60 stimuli which consisted of 10 repetitions of each of the six Cantonese test words produced by a native Cantonese speaker. Results show that among the six Cantonese tones, [55, 21, 25] are less difficult than [33, 22, 23] in both of production and perception. This is particularly for Cantonese tone [55] which can be correctly produced and perceived in all cases for all the three subjects. Cantonese tones [21] and [25] also can be correctly produced and perceived in most cases, although [21] cannot be correctly perceived by one subject and [25] cannot be correctly produced also by one subject. As for Cantonese tones [33, 22, 23], they are commonly mis-produced and mis-perceived by the subjects. In general, the subjects have similar performance in the production and perception of Cantonese tones and their performance in L2 (Cantonese) tones is affected by their native or L1 (Mandarin) tones. However, L1 transfer only happens in production rather than in perception.

1. INTRODUCTION

Tone languages are estimated to be spoken by over half of the world's population. Cantonese and Mandarin are tonal languages, in which tone plays an important role in distinguishing word meanings. Pike in his "Tone Language" (1976) describes that "A tone language may be defined as a language having lexically significant, contrastive, but relative pitch on each syllable". (p. 3)

In different tone languages, the number and type of tones greatly vary. Like Cantonese and Mandarin, though they are the two dialects of Chinese, their tone inventories differ. It has been considered that in Cantonese there are nine citation tones, including six long tones, i.e. three level tones [55, 33, 22], one falling tone [21], and two rising tones [25, 23], and three short level tones, i.e. [5, 3, 2] (Chao 1947, Zee 1999, Liu 2001, Flynn 2003). As for Mandarin, there are only four long citation tones, including the level tone [55], the falling tone [51], the rising tone [35], and the falling-rising tone [214] (Howie 1970, Chao 1976, Lin 1996, Lee and Zee 2003). Table 1 and Table 2 present the citation tones in Cantonese and those in Mandarin respectively. Also shown in the tables are the tonal categories and the tone values for the citation tones of the two dialects according to Zee (1999) and Lee and Zee (2003).

Register	*Ping* (level)	*Shang* (rising)	*Qu* (departing)	*Ru* (entering)	
Yin (Upper)	55	25	33	5	3
Yang (Lower)	21	23	22	2	

Table 1: Nine citation tones in Cantonese.

Register	*Ping* (level)	*Shang* (rising)	*Qu* (departing)
Yin (Upper)	55	214	51
Yang (Lower)	35		

Table 2: Four citation tones in Mandarin.

As can be seen in Table 1, in present-day Cantonese, all the eight historical tones in Chinese, i.e. the tones *yin ping* (upper level) [55], *yang ping* (lower level) [21], *yin shang* (upper rising) [25], *yang shang* (lower rising) [23], *yin qu* (upper departing) [33], *yang qu* (lower departing) [22], *yin ru* (upper entering) [5, 3] and *yang ru* (lower entering) [2] are retained. The six long tones [55, 21, 25, 23, 33, 22] in Cantonese are produced on open syllables ending in a vowel

or closed syllables with a final nasal consonant ([m], [n] or [ŋ]). As for the three short tones [5, 3, 2], they are only produced on checked syllables closed with a final stop consonant ([p], [t] or [k]). For Cantonese tone *yin ru* on checked syllables, it has been split into two as [5] and [3] at the present time according to the length of vowel in the syllables, with [5] associated with a short vowel and [3] associated with a long vowel. Phonologically, the three *ru* tones [5, 3, 2] are treated as the short variants of the three long tones [55, 33, 22] respectively due to the similarity in the pitch level and contour of the two sets of tones. It follows that there are only six contrastive tones, i.e. [55(5), 21, 25, 23, 33(3), 22(2)], in Cantonese.

As for Mandarin tones shown in Table 2, there are only four of the historical tones in Chinese retained in present-day Mandarin, together with a neutral tone treated as the fifth tone which is not shown in the table. Unlike Cantonese, tones in Mandarin have undergone a drastic reduction. The *yin-yang* (upper-lower) distinction is still retained for the *ping* tones, i.e. *yin ping* (upper level) and *yang ping* (lower level), however not for the *shang* and *qu* tones. There is only one *shang* (rising) tone and one *qu* (departing) tone in present-day Mandarin, and all the short *ru* (entering) tones have disappeared. Thus, all of the four Mandarin citation tones are long, and they are produced on open syllables ending in a vowel or closed syllables with a nasal ending ([m] or [n]).

A comparison of Table 1 and Table 2 shows that Cantonese has a richer tone inventory than Mandarin, due to the disappearance of the *ru* (entering) tones and the merge of the *yin* (upper) and *yang* (lower) registers for the *shang* (rising) and *qu* (departing) tones in Mandarin. The tone system of Mandarin is also less complicated than that of Cantonese due to striking similarity in pitch contour or shape between the tones in Cantonese rather than in Mandarin. For instance, there is only one level and one rising tones ([55] and [35]) in Mandarin, but there are three level ([55, 33, 22]) and two rising ([25, 23]) tones in Cantonese. This suggests that Cantonese tone system may be characterized as a register type and Mandarin tone system as a contour type. The distinctive feature of the tones in the contour system is pitch contour, such as rising, falling, dipping or level, while the distinctive feature of the tones in the register system is relative pitch level, such as high, mid or low, rather than contour shape. Table 3 shows the similarities and differences in contour and level of the tones between Cantonese and Mandarin.

Tone contour	Tone level	Cantonese	Mandarin
Level	High	[55, 5]	[55]
	Mid	[33, 3]	(none)
	Low	[22, 2]	(none]
Rising	High	[25]	[35]
	Low	[23]	(none)
Falling	High	(none)	[51]
	Low	[21]	(none)
Dipping (falling-rising)	Low	(none)	[214]

Table 3: Contour and level of the tones in Cantonese and Mandarin.

As shown in Table 3, except for the tone [55] which occurs in both Cantonese and Mandarin, all the other tones in the two dialects are different. Due to the difference in tone value between the tones in Cantonese and Mandarin, Mandarin-speaking learners of Cantonese are expected to have difficulties in acquisition of Cantonese tones and to make errors either in production or perception. This study is aiming at the performance in production and perception of Cantonese tones by Mandarin-speaking learners.

There have been a number of studies about the production and perception of Cantonese and Mandarin tones by either the native or non-native speakers. Leung (2008) was to find out i) which Mandarin tones were the most difficult in perception for Cantonese speakers; ii) among the speakers of Cantonese, Mandarin and English, which group of speakers would make more errors in the production of Mandarin tones; and iii) what is the frequent substitution made by Cantonese speakers for Mandarin tones due to L1 (Cantonese) transfer. Best's Perceptual Assimilation Model (1995) was adopted in Leung's study for the perceptual assimilation patterns of Mandarin and Cantonese tones for Cantonese speakers. Results showed that for, the four Mandarin tones, [35, 214] were the most difficult to be identified, followed by [51] and [55], by Cantonese speakers. In production task of Mandarin tones, it was observed that the performance of Mandarin speakers was the best, with the lowest deviation scores which correlated with the degree of error, followed by Cantonese speakers and English speakers in decreasing order. For the perception of Mandarin tones by Cantonese speakers, both Mandarin [35, 214] were mapped to Cantonese [25] and both Mandarin [51, 55] were mapped to Cantonese [55]. The results support Best's Perceptual Assimilation

Model (1995) that the perception of non-native tones will be assimilated to the native tonal categories rather than to an uncategorized speech sound or to a non-speech sound, indicating the effect of L1 transfer.

So (2005) was a perceptual study of Mandarin tones by two groups of non-native speakers, Cantonese and Japanese. The aim was to examine to what extent the phonological and phonetic influences of the prosodic system of L1 on the perception of the tones in L2. Based on the findings of previous studies (Kiriloff 1969, Miracle 1989, Shen 1989, Lee, Vakoch, & Wurm 1996, Gottfried & Suiter 1997) two hypotheses were tested in So's study. The first hypothesis was that the performance of Cantonese speakers would be better than Japanese speakers, as Cantonese is a tone language, but not Japanese, and Cantonese speakers have more linguistic experience than Japanese speakers in using pitch in their native language. The second hypothesis was that both groups of speakers had difficulties in perceiving two pairs of Mandarin tones [55] & [51] and [35] & [214], with similar pitch level or shape for the two tones in each pair. The first hypothesis was not confirmed by the results obtained in the study. The performance of Cantonese speakers was similar to that of Japanese speakers, indicating that the linguistic experience of Cantonese speakers had no advantage in perceiving Mandarin tones. Furthermore, both Cantonese and Japanese speakers made considerable amount of errors in the identification of the tones [55] & [51] and [35] & [214] in the two pairs, although Cantonese speakers made more noticeable errors than the Japanese speakers did. The results however confirmed the second hypothesis that the pairs of tones [55] & [51] and [35] & [214] are difficult to be identified.

Li (2004) was a perceptual study of the six Cantonese contrastive tones [55, 33, 22, 21, 25, 23] by native speakers of Cantonese. Three identification tests of level tones, rising tones and falling tone in Cantonese were conducted in the study. It was found that the subjects were easier to distinguish between the tones [55] and [33] than between the tones [33] and [22], due to larger difference in F0 between the tones in the first pair than in the second pair. This result supports Chao's (1947) and Vance's (1976) characterizations of the three Cantonese level tones as [55, 33, 22] rather than Hashimoto's (1972) characterization as [55, 44, 33]. It was also found that the distinction between the two rising tones were not easy, due to large similarity in the F0 contours of the two rising tones. Thus, it was concluded that F0 played an essential role in the perception of Cantonese tones.

In my knowledge, there is lack of studies about the production and perception of Cantonese tones by Mandarin speakers. In this research paper, production and perception of the six long tones of Cantonese by native Mandarin speakers are investigated. Three Mandarin speakers were asked to read out a minimal set of monosyllabic words associated with the six Cantonese long tones, like the monosyllables [si 55], [si 25], [si 33], [si 21], [si 23] and [si 22], i.e. 思, 史, 試, 時, 市 and 事 in Chinese character. The subjects were also asked to identify the words of the minimal set associated with the six different Cantonese tones produced by a native Cantonese speaker in a listening test. This study aims to find out pattern of errors made by Mandarin-speaking learners of Cantonese in both the production and perception aspects and to determine the relationship between production and perception of tones by second language learners. Comparison of the production and perceptual results may demonstrate whether the speakers have difficulties in perceiving and pronouncing the same tone.

2. METHODOLOGY

2.1. Subject

In this study, three female subjects participated in the production and perception tests of Cantonese tones. All subjects were native speakers of Mandarin from Mainland China, who were the second or third year undergraduates of the City University of Hong Kong and in their early twenties. They were born and raised in their home towns only using Mandarin and no any other dialects in their daily life. They did not learn Cantonese before coming to Hong Kong for tertiary education. In Hong Kong, they spoke Cantonese only to Cantonese speakers and in most of the time, they only used Mandarin. All subjects reported that they could understand majority of what a Cantonese speaker said and none of them had a history of hearing and speaking problems.

2.2. Test Materials

For both the production and perception tests of this study, the test material was a minimal set of six Cantonese monosyllabic words containing the same syllable [si], associated with one of the six Cantonese long tones, i.e. [55, 25, 33, 21, 23, 22] (Table 4). All the test words are meaningful and familiar to the subjects.

Cantonese Long Tones	Test Words
55	[si 55] 思 (to think)
25	[si 25] 史 (history)
33	[si 33] 試 (to try)
21	[si 21] 時 (time)
23	[si 23] 市 (city)
22	[si 22] 事 (event)

Table 4: Test monosyllabic words associated with one of the six Cantonese long tones [55, 25, 33, 21, 23, 22] used in this study.

2.3. Production task

The three Mandarin subjects were asked to read out a word list, on which five repetitions of each of the six test words in Chinese characters were in randomized order. The subjects read the word list at a normal rate of speech, producing the test words one by one in isolation. A total of 30 test tokens (6 test words x 5 repetitions) were recorded from a subject for the production task.

Digital recordings were performed in the sound-proof booth in the Phonetics Laboratory of the Department of Chinese, Linguistics and Translation at the City University of Hong Kong. The recorded speech data were analyzed for the fundamental frequency (F0) characteristics of the tones associated with the Cantonese test words, using the pitch synchronized method provided by the speech analysis software CSL (Computerized Speech Lab) 4500 of Kay Elemetrics of USA available in the Phonetic Laboratory.

2.4. Perception task

The three Mandarin subjects took part in a perception test to identify the six Cantonese long tones after the production task. The stimuli used for the perception were the six same Cantonese test words used for the production task from the natural speech of a female native speaker of Cantonese. Three repetitions of each of the test words were made by the Cantonese speaker, and only the best token of each word was selected and used as the stimulus.

Figure 1 shows the F0 contours of the six Cantonese long tones, [55, 33, 22, 21, 25, 23], on the selected tokens from the Cantonese speaker. For each tone, 11 data points were sampled from its F0 contour, consisting of the onset (0%) of the F0 contour as the first data point and subsequent data points at every 10% of the overall duration of the F0 contour. The F0 contours in the figure were drawn by connecting the 11 data points (not shown in the figure) for each of the tones.

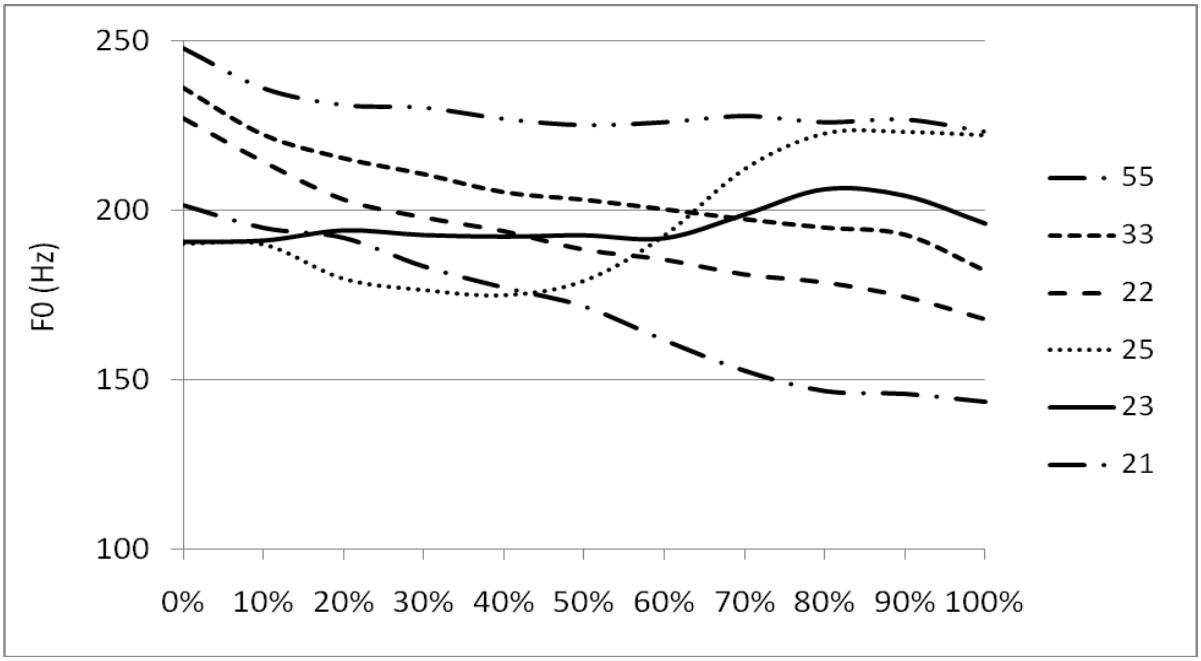


Figure 1: F0 contours of the six Cantonese tones [55, 33, 22, 21, 25, 23] from a female native speaker of Cantonese used as the stimuli for the perception test.

As shown in the figure, the six Cantonese tones are distinct from one another in pitch level and/or pitch shape of their F0 contours. The F0 contours of the tones [55, 33, 22] are basically level. While the F0 contours of these three tones tend to slightly falling toward the end of the tones, the degree of falling is much less pronounced than that of the F0 contour of the tone [21] which is being a falling tone in Cantonese. The pitch level (at the mid point) of the F0 contour is the highest for [55] (225 Hz), to be followed by [33] (202 Hz) and [22] (192 Hz) in a decreasing order. The pitch level of the F0 contour of the tone [55] reaches the upper level of the pitch range of the six tones for the speaker, and there is larger spacing or distance between the F0 contours of [55] and [33] (with the difference in 23 Hz) than between those of [33] and [22] (with the difference in 10 Hz). In spite of the small spacing between the F0 contours of [33] and [22], the F0 contours of these two tones are still distinct from one another.

As for the three Cantonese dynamic or contour tones for the speaker, the tone [21] is the only one which has a large drop in F0 (from 192 Hz at the mid point of the portion before falling to 151 Hz at the mid point of the portion after falling). The falling F0 contour of [21] is slightly lower in pitch than the F0 contour of [22] near the beginning, but the two tones become to have a large pitch difference toward the end of the tones due to a large drop in F0 for [21] to the lowest level of the pitch range of the six tones for the speaker. Similar to [21],

the F0 contours of [25] and [23] are lower in pitch than the F0 contour of [22] at the beginning and the F0 contours of the three tones become separate toward the end of the tones. Different from [21], both the F0 contours of [25] and [23] have an upward deflection after the first half of the tones. The upward deflection is much larger for [25] (from 175 Hz to 223 Hz), reaching the pitch level close to that of [55], than [23] (from 194 Hz to 207 Hz), reaching the pitch level similar to that of [33]. In general, the pitch patterns of the six Cantonese tones [55, 33, 22, 21, 25, 23] for the Cantonese speaker are similar to the characterizations of tones in the previous studies of Cantonese (Chao 1947, Zee 1999, Liu 2001, Flynn 2003) as well as the tone letters that have been assigned to the tones.

Ten repetitions of each selected token of the six Cantonese test words from the native Cantonese speaker making up a total of 60 stimuli (6 test words x 10 repetitions) were randomized in a block. The stimuli were played using the Window Media Player on a computer to the subjects through a headphone. An answer sheet containing the six Cantonese test words in Chinese character for each of the 60 trials was given to the subjects. The task of the subjects was to identify a Cantonese word that they heard in each trial by circling the corresponding Chinese character of the word on the answer sheet. There was no time limit for the subjects to make a response to each stimulus, but they were forced to response even they found difficulty in identifying the words. Thus, a total of 300 responses (5 subjects x 60 stimuli) were obtained for analyzing the correct and error patterns of the perception of the six Cantonese long tones by the Mandarin-speaking subjects.

3. RESULTS

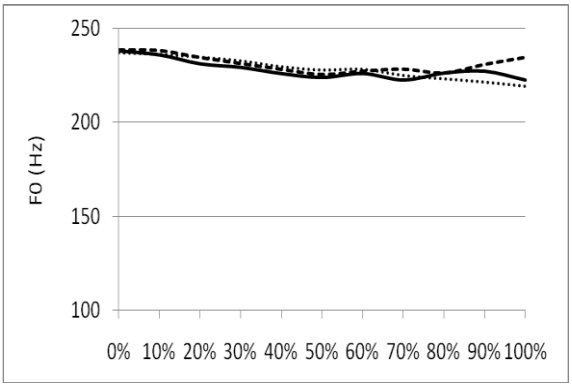
3.1. Production task

In this section, the results of the frequency analysis of the F0 contours of the six Cantonese long tones, i.e. [55, 33, 22, 21, 25, 23], produced by each of the three Mandarin speakers will be presented one by one. The patterns of the production of the six Cantonese tones for all the three speakers will be generalized afterward.

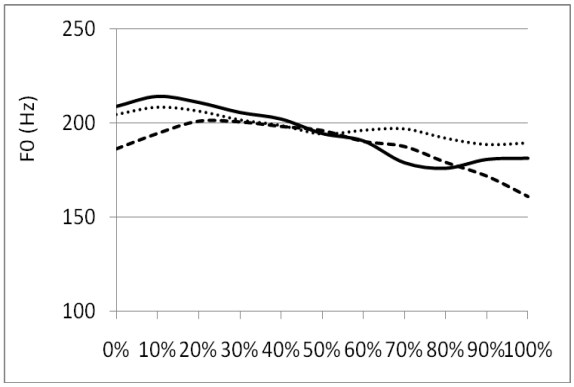
3.1.1. Mandarin Speaker 1

Figures 2(a)-(f) show the F0 contours of the six Cantonese long tones on the test words 思, 史, 試, 時, 市 and 事 for Mandarin Speaker 1. In each figure, the three superimposed F0 contours represent the three best tokens of a particular Cantonese long tone.

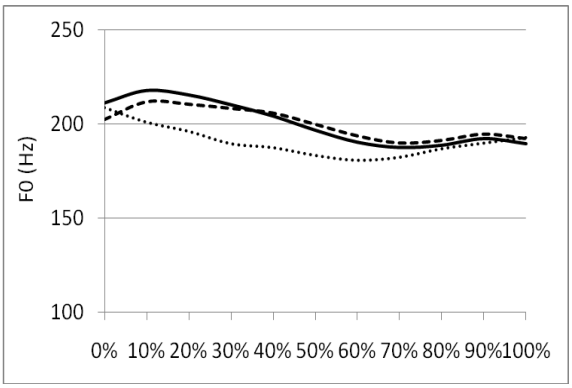
2(a) F0 contours of [si 55] 思



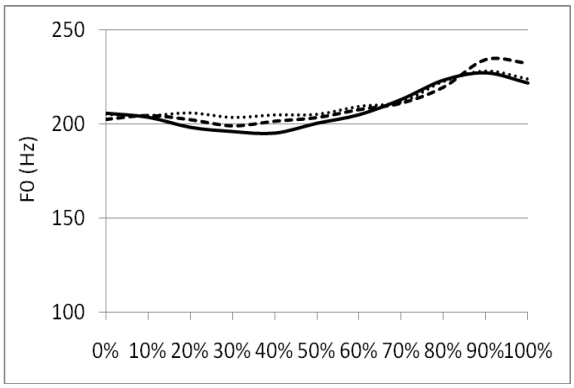
2(d) F0 contours of [si 21] 時



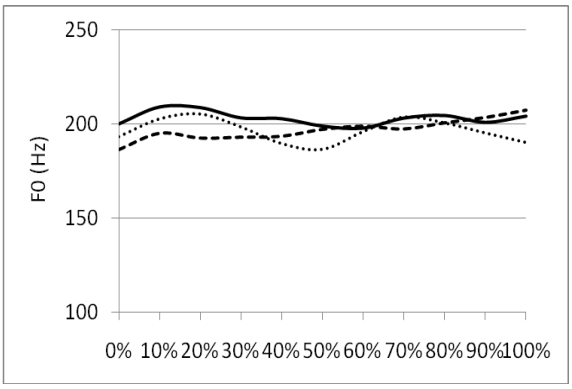
2(b) F0 contours of [si 33] 試



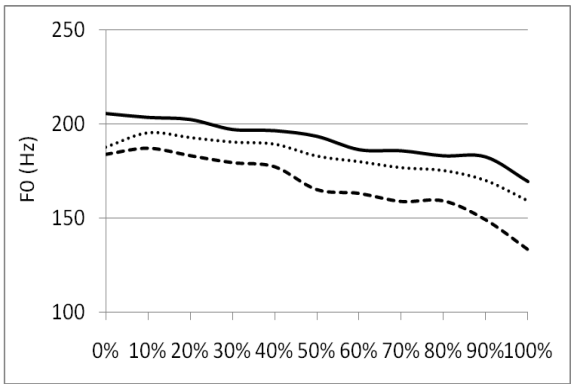
2(e) F0 contours of [si 25] 史



2(c) F0 contours of [si 22] 事



2(f) F0 contours of [si 23] 市



Figures 2(a)-(f): F0 contours of the six Cantonese long tones [55, 33, 22, 21, 25, 23] for Mandarin Speaker 1.

With regard to the three 'level' tones in Cantonese, i.e. [55, 33, 22], produced by Mandarin Speaker 1 (Figures 2(a)-(c)), [55] remains as a 'level' tone, whereas [33] becomes a 'falling' tone and [22] varies as a 'level' or 'dipping' tone. As shown in Figure 2(a), the three F0 contours of [55] are basically 'level', despite a small upward or downward deflection near the end of the tone. The mean frequency value taken at the mid point of the flat portion of the F0 contours of the three tokens of [55] is 220 Hz, which is the highest among the six Cantonese long tones produced by this subject. Thus, Cantonese tone [55] is correctly produced, remaining as a 'high level' tone in the speech of Mandarin Speaker 1.

As for Cantonese tone [33], it cannot be correctly produced as a 'level' tone by this speaker. As shown in Figure 2(b), the F0 contours of [33] are basically 'falling', while there is a small upward deflection near the end of the tone. The mean frequency value for the three F0 contours of [33] drops from 201 Hz (at the highest point of the F0 contours before falling) to 187 Hz (at the lowest point of the F0 contours after falling). Relative to the F0 values for the other Cantonese tones produced by Mandarin Speaker 1, [33] may be characterized as a 'mid to low-mid falling' tone.

As for the F0 contours of Cantonese tone [22] shown in Figure 2(c), two of them remain as 'level', with the mean F0 value of 198 Hz (averaging the values taken at the mid point of the flat portion of the two F0 contours), but the remaining one becomes 'falling-rising' with the F0 value decreasing from 205 Hz (at the point of the F0 contour before falling) to 186 Hz (at the turning point of the F0 contour), and then returning to 204 Hz (at the highest point of the

F0 contour after rising). Thus, [22] for Mandarin Speaker 1 may be characterized as a ‘mid level’ or ‘mid to low to mid falling-rising’ tone.

As for the three Cantonese dynamic tones [21, 25, 23] for this speaker, [21] and [25] remain as a ‘falling’ and a ‘rising’ tones respectively, but [23] is wrongly pronounced as a ‘falling’ tone.

As shown in Figure 2(d), the three F0 contours of the tone [21] are falling in various degrees. On average, the mean F0 value of [21] drops from 208 Hz (at the highest point of the F0 contours before falling) to 177 Hz (at the lowest point of the F0 contour after falling) which is close to the lowest level of the pitch range of all the six tones produced by the speaker. In general, the tone [21] is appropriately produced, while it may be characterized as a ‘mid to low falling’ tone.

Cantonese tone [25] is also appropriately produced by this speaker. As shown in Figure 2(e), the three F0 contours of [25] have an upward deflection after the first half of the tone. On average, the tone is increasing from the mid level, with the F0 value of 202 Hz (at the mid point of the first half portion of the F0 contours before rising), to 226 Hz (at the highest point of the F0 contours after rising), which is close to the pitch level of the tone [55] (Figure 2(a)) produced by the speaker. Thus, [25] for Mandarin Speaker 1 remains as a ‘mid to high rising’ tone.

As for the other rising tone [23] in Cantonese, it changes to a ‘falling’ tone in the speech of Mandarin Speaker 1. As can be seen in Figure 2(f), for all the three tokens of [23] produced by this speaker have a large drop in F0 from the mean value of 195 Hz (at the highest point of the F0 contours before falling) to 154 Hz (at the lowest point of the F0 contours after falling). The F0 contours and levels of the tone [23] are similar to those of the tone [21] (Figure 2(d)) produced by this speaker. Thus, for Mandarin Speaker 1, the tone [23] merges with the tone [21] to become a ‘mid to low falling’ tone.

The F0 characteristics of the six long Cantonese tones produced by Mandarin Speaker 1 are summarized in Table 5. As can be seen, of the six tones, only [55], [21], and [25] are appropriately produced by the speaker. As for the other three tones, [22] may vary as a ‘mid level’ or ‘mid to low to mid falling-rising’ tone; and both of [33] and [23] become ‘falling’ as

a ‘mid to low-mid falling’ tone and a ‘mid to low falling’ tone respectively. Since [23] becomes the same as [21], there are only five tones still distinguished in the speech of Mandarin Speaker 1.

Cantonese tones	Realizations in the speech for Mandarin Speaker 1
[55]	Remains as a ‘high level’ tone
[33]	Becomes a ‘mid to low-mid falling’ tone
[22]	Becomes a ‘mid level’ or ‘mid to low to mid falling-rising’ tone
[21]	Remains as a ‘mid to low falling’ tone
[25]	Remains as a ‘mid to high rising’ tone
[23]	Becomes a ‘mid to low falling’ tone

Table 5: A summary of the F0 characteristics of the six Cantonese long tones produced by Mandarin Speaker 1.

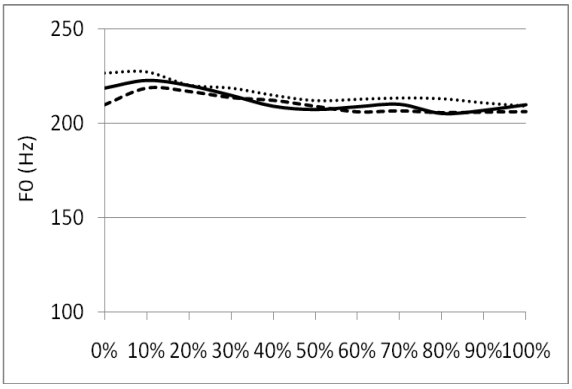
3.1.2. Mandarin Speaker 2

The F0 contours of the six Cantonese long tones produced by Mandarin Speaker 2 are presented in Figure 3(a)-(f). For this speaker, similar to Mandarin Speaker 1, the tone [55] is correctly produced as ‘high level’. As shown in Figure 3(a), the three F0 contours of [55] basically remain constant throughout the tone, with the mean F0 value of 209 Hz for the three tokens which is the highest among the six tones produced by the speaker. As for the other two ‘level’ tones [33] and [22] in Cantonese, [33] becomes ‘falling’ (Figure 3(b)) and [22] varies as ‘level’ or ‘falling’ (Figure 3(c)) in the speech of Mandarin Speaker 2. As shown in Figure 3(b), all the three F0 contours of [33] are ‘falling’, with the mean F0 value dropping from 201 Hz to 176 Hz. As for [22] presented in Figure 3(c), two tokens are realized as a ‘falling’ tone, with the mean F0 value dropping from 198 Hz to 163 Hz, whereas the remaining one is realized as a ‘level’ tone, with the mean F0 value of 174 Hz.

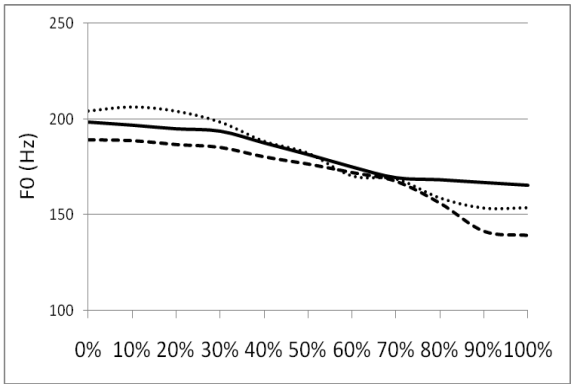
As for the three Cantonese dynamic tones produced by this speaker, the tone [21] is correctly produced as ‘falling’, with the mean F0 value decreasing from 197 Hz to 153 Hz. As for the two ‘rising’ tones [25] and [23] in Cantonese, both of them become a dipping ‘falling-rising’ tone and have similar F0 value (Figures 3(e) and 3(f)). The mean F0 value for the tone [25]

falls from 185 Hz to 163 Hz, and then returns to 197 Hz. As for the tone [23], the mean F0 value drops from 202 Hz to 170 Hz and then returns to 188 Hz.

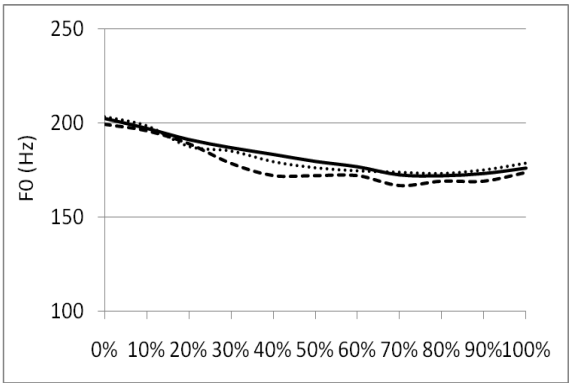
3(a) F0 contours of [si 55] 思



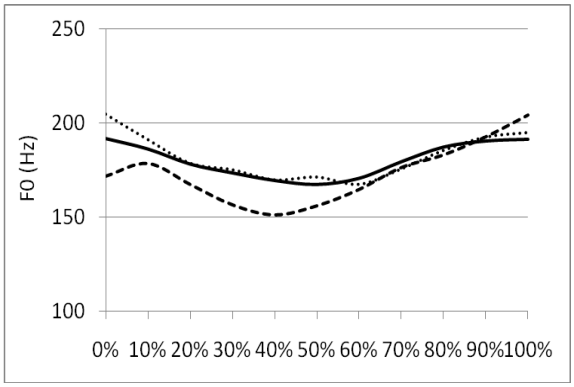
3(d) F0 contours of [si 21] 時



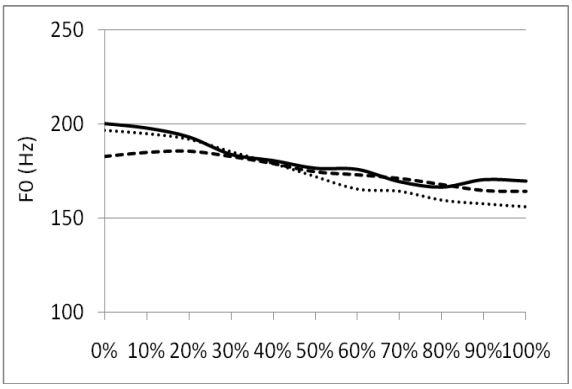
3(b) F0 contours of [si 33] 試



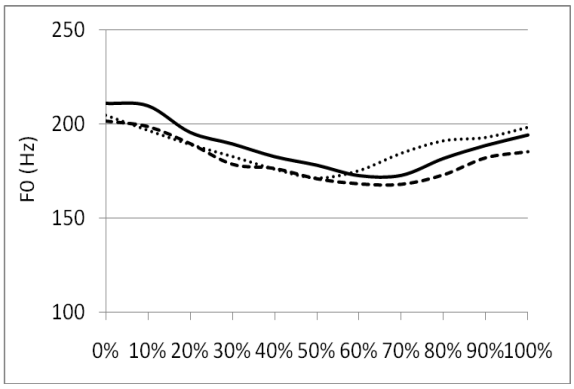
3(e) F0 contours of [si 25] 史



3(c) F0 contours of [si 22] 事



3(f) F0 contours of [si 23] 市



Figures 3(a)-(f): F0 contours of the six Cantonese long tones [55, 33, 22, 21, 25, 23] for Mandarin Speaker 2.

As summarized in Table 6, in general, the six Cantonese long tones for Mandarin Speaker 2 are divided into four groups, i.e. [55], [33, 22, 21], [22] and [25, 23], and they may be characterized as the ‘high level’, ‘mid to low falling’, ‘mid level’ and ‘mid to low to mid falling-rising’ tones respectively. For this speaker, only the tones [55] and [21] are correctly produced, whereas the tone [33] merges with the tone [21], the tone [22] may become a ‘mid level’ or ‘mid to low falling’ tone and both the tones [25, 23] become the same ‘dipping’ tone. As a result, there are three or four tones that are still distinct in the speech of Mandarin Speaker 2.

Cantonese tones	Realizations in the speech for Mandarin Speaker 2
[55]	Remains as a ‘high level’ tone
[33]	Becomes a ‘mid to low falling’ tone
[22]	Becomes a ‘mid level’ or ‘mid to low falling’ tone
[21]	Remains as a ‘mid to low falling’ tone
[25]	Becomes a ‘mid to low to mid falling-rising’ tone
[23]	Becomes a ‘mid to low to mid falling-rising’ tone

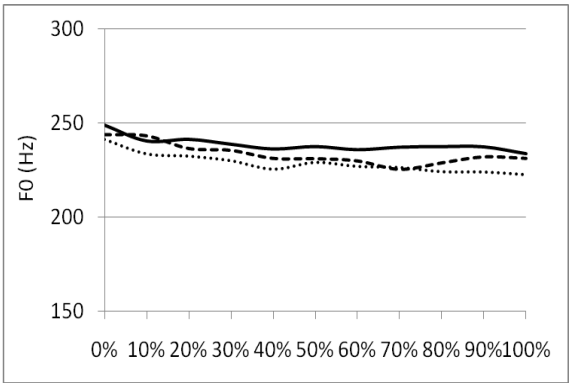
Table 6: A summary of the F0 characteristics of the six Cantonese long tones produced by Mandarin Speaker 2.

3.1.3. Mandarin Speaker 3

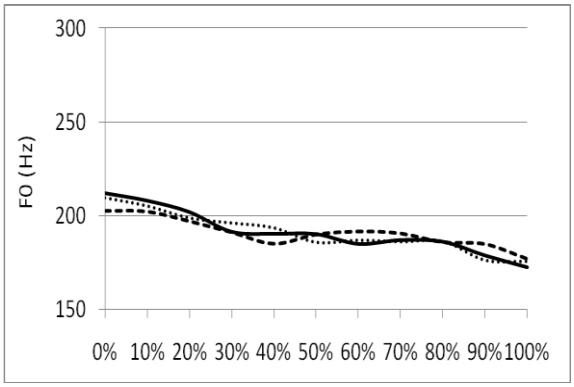
Figures 4(a)-4(f) show the F0 contours of the six Cantonese long tones for Mandarin Speaker 3. For the tone [55] produced by this speaker (Figure 4(a)), it is similar to the tone [55] for Mandarin Speakers 1 and 2 (Figures 2(a) and 3(a)) that it remains as a ‘high level’ tone. The mean F0 value for the three tokens of [55] is 231 Hz which is the highest among the six tones produced by the speaker. As for the other two ‘level’ tones in Cantonese, the three F0 contours of [33] for Mandarin Speaker 3 shown in Figure 4(b) are ‘falling’ instead of the appropriate ‘level’ shape, with the mean F0 value dropping from 221 Hz to 175 Hz. As for the tone [22] for this speaker, it is produced correctly as a ‘level’ tone in one token, but becomes a ‘falling’ tone in the remaining two tokens. For the ‘level’ token, the mean F0 value is 204 Hz, whereas for the two ‘falling’ tokens, the mean F0 value is decreasing from 216 Hz to 182 Hz.

As for the three dynamic Cantonese tones [21, 25, 23] produced by Mandarin Speaker 3, both [21] and [25] are correctly produced as a ‘falling’ tone and a ‘rising’ tone respectively (Figures 4(d) and 4(e)), but [23] is wrongly produced as a ‘falling’ tone (Figure 4(f)). The mean F0 value is decreasing from 208 Hz to 175 Hz for [21], increasing from 201 Hz to 247 Hz for [25], and decreasing from 211 Hz to 177 Hz for [23].

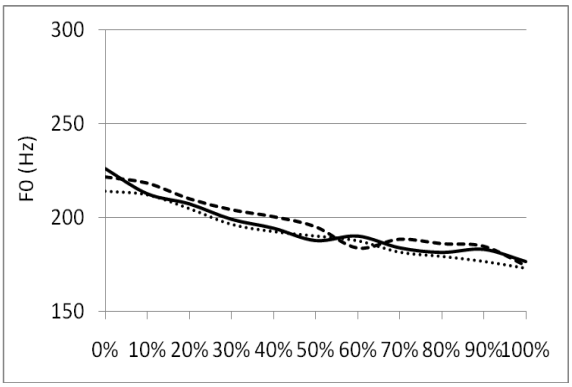
4(a) F0 contours of [si 55] 思



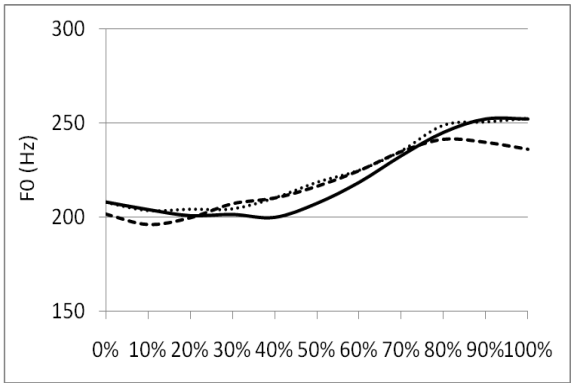
4(d) F0 contours of [si 21] 時



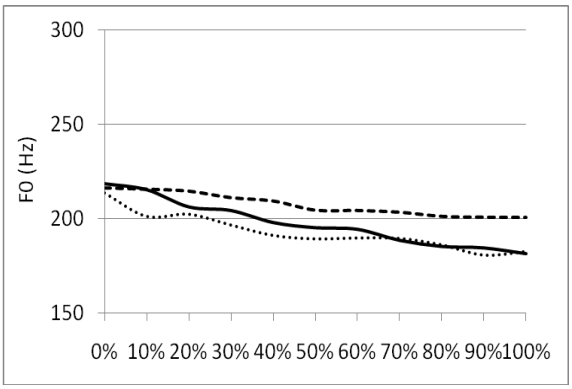
4(b) F0 contours of [si 33] 試



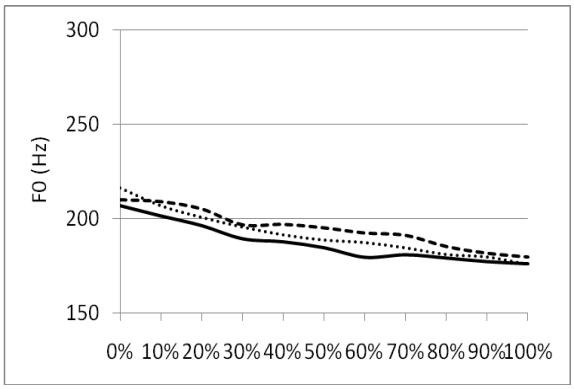
4(e) F0 contours of [si 25] 史



4(c) F0 contours of [si 22] 事



4(f) F0 contours of [si 23] 市



Figures 4(a)-(f): F0 contours of the six Cantonese long tones [55, 33, 22, 21, 25, 23] for Mandarin Speaker 3.

As summarized in Table 7, in general, [55] is produced as a ‘high level’ tone, [33] a ‘mid to low falling’ tone, [22] a ‘mid to low falling’ or ‘low-mid level’ tone, [21] a ‘mid to low falling’ tone, [25] a ‘low-mid to rising’ tone, and [23] as ‘mid to low falling’ tone for Mandarin Speaker 3. Since the tones [33, 23] and also the tone [22] may become the same as the tone [21], there are only three or four tones that are still distinct in the speech for this speaker.

Cantonese tones	Realizations in the speech for Mandarin Speaker 3
[55]	Remains as a ‘high level’ tone
[33]	Becomes a ‘mid to low falling’ tone
[22]	Remains as a ‘low-mid level’ tone or becomes a ‘mid to low falling’ tone
[21]	Remains as a ‘mid to low falling’ tone
[25]	Remains as a ‘low-mid to high rising’ tone
[23]	Becomes a ‘mid to low falling’ tone

Table 7: A summary of the F0 characteristics of the six Cantonese long tones produced by Mandarin Speaker 3.

Let’s draw conclusions about the performance of the three Mandarin subjects did in the production task. A comparison of the F0 characteristics of the six Cantonese long tones produced by the three subjects as summarized in Tables 5-7 shows that all the subjects can produce the tone [55] appropriately as a ‘high level’ tone. For this, it is not surprising as the ‘high level’ tone [55] occurs in both Cantonese and Mandarin tonal systems. Furthermore, the test word 思 associated with Cantonese tone [55] used for the investigation is also produced with the same tone [55] in Mandarin. For this case, the L1 tonal knowledge of the speakers may contribute to the production of the non-native tone.

Besides the tone [55], there is a similar ‘rising’ tone in Cantonese and Mandarin, which may be characterized as a ‘mid-low to high rising’ [25] or a ‘mid to high rising’ [35] respectively in the two languages. So, it is expected that this ‘rising’ tone should also be correctly produced by Mandarin subjects. However, as can be seen in Tables 5-7, only for Mandarin

Speakers 1 (Table 5) and 3 (Table 7) Cantonese tone [25] remains as a ‘high rising’ tone, while in the speech of these two speakers the tone may vary slightly as ‘mid to high rising’ (Mandarin Speaker 1) or ‘low-mid to high rising’ (Mandarin Speaker 3). As for Mandarin Speaker 2, Cantonese [25] is wrongly produced as a ‘falling-rising’ dipping tone. The performance of this speaker may be due to the effect of L1 transfer, where the test word 史 associated with Cantonese [25] used for the analysis is produced with ‘falling-rising’ dipping tone [214] in Mandarin.

As for the tones [33, 22, 21, 23] in Cantonese, all of them are not occurring in Mandarin. For these four tones in Cantonese, it is surprising that [21] generally remains unchanged as ‘mid to low falling’ in the speech of all the three Mandarin speakers, although the degree of falling of this tone for the Mandarin speakers (Figures 2(d), 3(d) and 4(d)) is less pronounced as that of the tone [21] produced by the native Cantonese speaker in this study (Figure 1). For this case, it may be also related to the native or L1 tones of speakers, as in Mandarin the ‘falling-rising’ dipping tone [214] has a ‘low falling’ variant from [21].

As for the other three tones in Cantonese which are not occurring in Mandarin, interestingly, Cantonese tone [22] may remain as ‘level’ in some tokens for all the three Mandarin speakers, while it is slightly varied among the three speakers as a ‘mid level’ tone (Mandarin Speakers 1 and 2) or a ‘low-mid level’ tone (Mandarin Speaker 3). In the other tokens of this tone produced by the three Mandarin speakers, it may become a ‘mid to low to mid falling-rising’ tone (Mandarin Speaker 1) or a ‘mid to low falling’ tone (Mandarin Speakers 2 and 3). For the latter case, it may be due to the fact that the test word 事 associated with Cantonese tone [22] used for this study is produced with the ‘falling’ tone [51] in Mandarin. Among the six Cantonese long tones produced by the Mandarin speakers, both the inter-variation between the speakers and the intra-variation between the tokens for the same speaker are the largest for the tone [22].

As for Cantonese tones [33] and [23] which are also non-occurring in Mandarin, both of them are wrongly produced by the three Mandarin speakers. For [33], it becomes a ‘falling’ tone, as a ‘mid to low-mid falling’ tone for Mandarin Speaker 1 or a ‘mid to low falling’ tone for Mandarin Speakers 2 and 3. As for [23], it also becomes a ‘mid to low falling’ tone for Mandarin Speakers 1 and 3, while it is a ‘falling-rising’ dipping tone for Mandarin Speaker 2. Generally speaking, Cantonese tones [33] and [23] tend to become neutralized in the speech

of the Mandarin speakers. Concerning the change of the tones [33] and [23] to ‘falling’, it may be due to the fact that the test words 試 and 市 used for the respective Cantonese tones [33] and [23] are produced with the ‘high falling’ tone [51] in Mandarin.

On the whole, the data on the production of the six Cantonese tones, i.e. [55, 33, 22, 21, 25, 23], by the Mandarin speakers indicate the effect from the native or L1 tones of the speakers. Among the six Cantonese tones, [33, 22, 23] which are not occurring in Mandarin are the most difficult for the Mandarin speakers. For these tones, the most frequent substitution is similar to the L1 tones associated with the test words. The less difficult Cantonese tones for the three Mandarin speakers are [55, 25, 21], for which Cantonese [55] is the same as Mandarin [55], Cantonese [25] is similar to Mandarin [35], and Cantonese [21] is the same as to the variant from [21] of Mandarin [214].

3.2. Perception task

The results of the three subjects in the perception task will be discussed in this section. The responses (in absolute number and percentage) from each of the three speakers to the 60 stimuli consisting of the six Cantonese tones in 10 repetitions are presented in tables. The data of the three subjects will be compared and generalized for patterns of perception of the six Cantonese tones for Mandarin-speaking L2 Cantonese learners.

3.2.1. Mandarin Speaker 1

Table 8 presents the number and percentage of responses to the perception of the six Cantonese tones [55, 33, 22, 21, 25, 23] for Mandarin Speaker 1. The highlighted and underlined cases shown in diagonal of the table are correctly perceived by the subject. As can be seen from the table, this subject has no or the least difficulty in perceiving the three Cantonese tones [55, 21, 25]. For [55] and [21], they are correctly perceived in all cases (100%), and in a large majority of the cases (80%) [25] is also correctly perceived. As for the other three Cantonese tones [33, 22, 23], only in 20% of the cases they can be correctly perceived. For [33], most frequently (80% of the cases), it is misperceived as the ‘high level’ tone [55]. As for [23], in 60% of the cases it is wrongly perceived as the other ‘rising’ tone [25] in Cantonese. As for [22], it may be misperceived as [33], [21] or [25], and in each of the cases the perception rate is below 50%. Thus, for this subject, while it can be found that [22]

is never perceived as the ‘high level’ tone [55] or the ‘low rising’ tone [23], no pattern can be generalized for the misperception of [22] as [33], [21] or [25].

Stimuli / Responses	[55]	[33]	[22]	[21]	[25]	[23]
[55]	**10(100%)**	8(80%)	0	0	0	0
[33]	0	**2(20%)**	3(30%)	0	0	0
[22]	0	0	**2(20%)**	0	1(10%)	1(10%)
[21]	0	0	1(10%)	**10(100%)**	0	1(10%)
[25]	0	0	4(40%)	0	**8(80%)**	6(60%)
[23]	0	0	0	0	1(10%)	**2(20%)**

Table 8: Responses (in number and %) to the perception of the six Cantonese tones for Mandarin Speaker 1.

A summary of the perception data on the six Cantonese tones for Mandarin Speaker 1 is shown in Table 9. A comparison of the perception data presented in Table 9 and the production data presented in Table 5 for this speaker shows that the tones which can be correctly produced, i.e. [55, 21, 25], can also be correctly perceived, whereas the tones which are wrongly produced, i.e. [33, 22, 23], are also wrongly perceived. However, for the difficult tones [33, 22, 23] for this speaker, there is a variation between the performance in production and perception. In the production task, [33, 22, 23] are wrongly produced as a ‘mid to low-mid falling’ tone, a ‘mid to low to mid falling-rising’ tone, and a ‘mid to low falling’ tone respectively (Table 5), whereas in the perception task, these three tones are misperceived as the respective [55], [33/21/25] and [25] (Table 9).

Cantonese tones	Perception results for Mandarin Speaker 1
[55]	Correctly perceived as [55] in all cases
[33]	Misperceived as [55] in most cases
[22]	Perceived as [33, 22, 21, 25]
[21]	Correctly perceived as [21] in all cases
[25]	Correctly perceived as [25] in most cases
[23]	Misperceived as [25] in most cases

Table 9: A summary of the perception results of the six Cantonese tones for Mandarin Speaker 1.

3.2.2. Mandarin Speaker 2

As for Mandarin Speaker 2, the perception data on the six Cantonese tones are presented in Table 10. As shown in the table, similar to Mandarin Speaker 1, Mandarin Speaker 2 also has no or the least difficulty in perceiving Cantonese tones [55, 21, 25]. For these three tones, Mandarin Speaker 2 is similar to Mandarin Speaker 1 that both [55] and [21] are correctly perceived in all the cases (100%), and [25] is correctly perceived in most of the cases (70%). As for the other three Cantonese tones [33, 22, 23], Mandarin Speaker 2 is different from Mandarin Speaker 1 that both [33] and [22] can be correctly perceived in a half of the cases (50%). However, [33] may also be perceived as [22] or [23] (20% of the cases) or [55] (10% of the cases), and [22] may be perceived as [33] (30% of the cases) or [23] (20% of the cases). Concerning the tone [23], it is the most difficult one to be perceived by Mandarin Speaker 2, with the correct perception rate of 10% only. In the other cases, 40% of which [23] is misperceived as [22], 30% as [25], and 20% as [33]. As a result, no pattern can be generalized for the perception of [23] for this subject.

Stimuli / Responses	[55]	[33]	[22]	[21]	[25]	[23]
[55]	**10(100%)**	1(10%)	0	0	0	0
[33]	0	**5(50%)**	3(30%)	0	0	2(20%)
[22]	0	2(20%)	**5(50%)**	0	1(10%)	4(40%)
[21]	0	0	0	**10(100%)**	0	0
[25]	0	0	0	0	**7(70%)**	3(30%)
[23]	0	2(20%)	2(20%)	0	2(20%)	**1(10%)**

Table 10: Responses (in number and %) to the perception of the six Cantonese tones for Mandarin Speaker 2.

The results of perception of the six Cantonese tones for Mandarin Speaker 2 are summarized in Table 11. A comparison of the perception data in Table 11 and the production data presented in Table 6 for this speaker shows that the two Cantonese tones [55, 21] which can be 100% perceived are also correctly produced, which is the same as the case of Mandarin Speaker 1 as presented before. As for the other four Cantonese tones [33, 22, 25, 23] which are wrongly produced by Mandarin Speaker 2 (Table 6), they can be correctly perceived in most of the cases, except for [23] (Table 11). Furthermore, in the production task, [33, 22] may become the same ‘mid to low falling’ tones and [25, 23] become the same ‘mid to low to mid falling-rising’ tones. However, [33] and [22] as well as [25] and [23] are not confused with each other in perception. On the whole, the production and perception data for Mandarin Speaker 2 indicate that for the six Cantonese tones, [55, 21] are the least difficult tones in both production and perception; [33, 22, 25] are less difficult in perception than production; and [23] is the most difficult tone in both production and perception. The data for this speaker also show that the tones which can be correctly produced can also be correctly perceived (such as [55] and [21]), but not true in vice versa (such as [33, 22, 25]). It follows that the performance of this speaker is better in tone perception than in tone production.

Cantonese tones	Perception results for Mandarin Speaker 2
[55]	Correctly perceived as [55] in all cases
[33]	Perceived as [33] (more frequent) and [55, 22, 23,]
[22]	Perceived as [22] (more frequent) and [33, 23]
[21]	Correctly perceived as [21] in all cases
[25]	Correctly perceived as [25] in most cases
[23]	Perceived as [33, 22, 25, 23]

Table 11: A summary of the perception results of the six Cantonese tones for Mandarin Speaker 2.

3.2.3. Mandarin Speaker 3

Stimuli / Responses	[55]	[33]	[22]	[21]	[25]	[23]
[55]	**10(100%)**	9(90%)	1(10%)	0	0	0
[33]	0	**1(10%)**	2(20%)	3(30%)	0	2(20%)
[22]	0	0	**4(40%)**	4(40%)	0	0
[21]	0	0	2(20%)	**2(20%)**	3(30%)	6(60%)
[25]	0	0	0	0	**7(70%)**	1(10%)
[23]	0	0	1(10%)	1(10%)	0	**1(10%)**

Table 12: Responses (in number and %) to the perception of the six Cantonese tones for Mandarin Speaker 3.

As for Mandarin Speaker 3, the perception data on the six Cantonese tones [55, 33, 22, 21, 25, 23] presented in Table 12 show that [55] and [25] are correctly perceived in all (100%) or most (70%) of the cases, which is the same as the perception performance of these two tones

for Mandarin Speakers 1 and 2. However, as for [21], while it is 100% correctly perceived by Mandarin Speakers 1 and 2, it is correctly perceived only in 20% of the cases for Mandarin Speaker 3. In the other cases, [21] is misperceived as [33] (30% of the cases), [22] (40% of the cases), and [23] (10% of the cases). As for the tones [33, 22, 23], they are also difficult for Mandarin Speaker 3. In a majority of the cases, [33] is misperceived as [55] (90% of the cases) and [23] is misperceived as [21] (60% of the cases). As for [22], while in 40% of the cases it is correctly perceived as [22], it may be misperceived as [33] or [21] in 20% of the cases and misperceived as [55] or [23] in 10% of the cases.

Therefore, as summarized in Table 13, for Mandarin Speaker 3, only the tones [55] and [25] can be correctly perceived. The tones [33] and [23] are misperceived as [55] and [21] respectively. As for the perception of the tones [22] and [21], no pattern can be generalized, where [22] may be perceived as any tone, except for [25], and [21] may be perceived as [33, 22, 21, 23].

Cantonese tones	Perception results for Mandarin Speaker 3
[55]	Correctly perceived as [55] in most cases
[33]	Misperceived as [55] in most cases
[22]	Perceived as [55, 33, 22, 21, 23]
[21]	Perceived as [33, 22, 21, 23]
[25]	Correctly perceived as [25] in most cases
[23]	Misperceived as [21] in most cases

Table 13: A summary of the perception results of the six Cantonese tones for Mandarin Speaker 3.

A comparison of the perception data presented in Table 13 and production data in Table 7 for Mandarin Speaker 3 shows that only two tones, [55] and [25], can be correctly perceived and these two tones can also be appropriately produced. As for the other four tones, [33, 22, 21, 23], that cannot be correctly perceived, they are also wrongly produced, except for [21]. Thus, for this speaker, among the six tones, [55] and [25] are the least difficult tones in both production and perception, followed by [21] which is not difficult in production and only in perception, and [33, 22, 23] are the most difficult tones in both production and perception.

The production and perception data for this speaker also show that the tones which can be correctly perceived can also correctly produced (such as [55, 25]), but not true in vice versa (such as [21]). As for the tones which cannot be correctly produced (such as [33, 22, 23]), they are also difficult to be perceived. It follows that the performance of this speaker is better in tone production than in tone perception.

Overall, the perception data for the three Mandarin speakers (Tables 9, 11 and 13) show that the perception of Cantonese tones [55] and [25] are not difficult for all the subjects. It is followed by Cantonese tone [21], which can be correctly perceived by two of the three subjects (Mandarin Speakers 1 and 2), and Cantonese tones [33, 22, 23] are the most difficult tones in perception, although [33] and [22] are less frequent to be misperceived by one subject (Mandarin Speaker 2). Generally speaking, the perception data for the three subjects correspond to their production data as presented in Section 3.1. In both production and perception of Cantonese tones, [55, 25, 21] are less difficult than [33, 22, 23] for the Mandarin subjects. It follows that the tones which can be correctly produced can also be correctly perceived, whereas the tones which cannot be correctly produced are also difficult to be perceived appropriately. And, for each case, it is true in vice versa. It is considered that the performance in the production and perception of Cantonese tones for Mandarin subjects is affected by the native or L1 tones of the subjects. As noted before, Cantonese [55] is the same as Mandarin [55], Cantonese [25] is similar to Mandarin [35], and Cantonese [21] is the same as the variant form [21] of Mandarin [214]. Thus, it may be the reason for why Cantonese tones [55, 25, 21] are less difficult for the Mandarin subjects in both production and perception. As for Cantonese tones [33, 22, 23], they are not occurring in Mandarin and they are the difficult tones for the Mandarin subjects.

4. CONCLUSION

This paper has presented the results of both the production and perception of the six Cantonese long tones [55, 33, 22, 21, 25, 23] by the three Mandarin-speaking learners. The findings have demonstrated that the performance in both the production and perception of the tones in L2 is affected by the native or L1 tones of the subjects. The L2 tones which are similar to those in L1 (i.e. Cantonese tones [55, 25, 21]) are less difficult than L2 tones which are not occurring in L1 (i.e. Cantonese tones [33, 22, 23]). In general, the tones which are difficult to be produced are also difficult to be perceived and it is also true in vice versa. However, it should be noted that for the tones which are wrongly produced by the subjects, they may become similar to the native tones of the subjects. As for the tones which are wrongly perceived by the subjects, they are confused with the other tones in L2. Thus, it may be considered that the effect of L1 on L2 is more in production than perception.

Since the findings of this study are based on the performance of three Mandarin subjects only and there are some variations between the speakers, data from more subjects may be obtained to confirm the patterns observed in this study. Furthermore, more test words, where the tones associated with the test words may be similar or different in L1 on L2 of the subjects, can be used for further investigation of the effect of L1 transfer on the acquisition of L2 tones.

Nevertheless, the results of this study still provide useful information for Mandarin-speaking learners of Cantonese or teachers who teach Mandarin speakers Cantonese, for instance Cantonese tones [33, 22, 23] should receive more attention and need to have more training, either in production or perception, than the other tones.

5. ACKNOWLEDGEMENTS

I would like to make use of this opportunity to thank Dr Lee Wai Sum for her comments and suggestions.

6. REFERENCES

Chao, Yuen Ren. (1947). *Cantonese Primer*. Cambridge, Mass: Harvard University Press.

Chao, Yuen Ren. (1976). *Mandarin Primer*. Cambridge, Mass: Harvard University Press.

Flynn, Choi Yueng Chang. (2003). *Intonation in Cantonese*. Muenchen: LINCOM Europa.

Fok Chan, Yuen-Yuen. (1974). *A Perceptual Study of Tones in Cantonese*. Centre of Asian Studies, University of Hong Kong Press.

Howie, John Marshall. (1970). *The Vowel and Tones of Mandarin Chinese*. Retrieved from http://0-pqdd.sinica.edu.tw.lib.cityu.edu.hk/twdaoeng/servlet/advanced?query=7112454.

Lee, Wai Sum and Eric Zee (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, Vol 33, No.1, p.109-112.

Leung, Anna S. (2008). *Tonal Assimilation Patterns of Cantonese L2 Speakers of Mandarin in the Perception and Production of Mandarin Tones.* Ph.D. Dissertation, University of Toronto.

Li, Po Yee, (2004). *Perceptual Analysis of the Six Contrastive Tones in Cantonese.* International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages, Beijing, China, March 28-31, 2004.

Lin, Hua (1996). *Mandarin Tonology*. Taipei: Pyramid Press.

Liu, J. (2001). *Tonal Behavior in Some Tone Languages*. Ph.D. Dissertation, City University of Hong Kong.

Pike, Kenneth L. (1976). *Tone Languages*. Ann Arbor: The University of Michigan Press.

So, Connie K. (2005). *Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences.* Ph.D. Dissertation, University of Western Sydney.

Zee, Eric, 1999. Chinese (Hong Kong Cantonese). In the *Handbook of the International Phonetic Association*, p.58-60. Cambridge: Cambridge University Press.