Title	Tone discrimination of Cantonese and Mandarin between tonal and non-tonal speakers
Author(s)	Chan, Hol-yan, Cherry; 陳可人
Citation	Chan, H. C. [陳可人]. (2013). Tone discrimination of Cantonese and Mandarin between tonal and non-tonal speakers. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
Issued Date	2013
URL	http://hdl.handle.net/10722/238552
Rights	The author retains all proprietary rights, (such as patent rights) and the right to use in future works.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Tone Discrimination of Cantonese and Mandarin between Tonal and Non-Tonal Speakers

Cherry Chan Hol Yan

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2013.

Tone Discrimination of Cantonese and Mandarin between Tonal and Non-Tonal Speakers

Chan Hol Yan Cherry

Abstract

This study investigated the perception of Cantonese and Mandarin tones by both tonal and non-tonal language speakers. Eighty subjects were recruited from four language groups (native Cantonese, Mandarin, English and Cantonese listeners who are Mandarin naïve). There were two parts in this study: the first experiment focused on Cantonese tones, while the second experiment focused on Mandarin tones. Participants were asked to discriminate if there were any tone differences between two target words in carrier phrase. Results showed that linguistic background may not necessarily influence the perception of non-native tones, and if influence was present, it could be positive or negative. Erroneous tone patterns from different native language groups were compared in regard to their native language's tone inventory and also in terms of psychoacoustic features. Other factors that may influence cross-language tone perceptions were also discussed.

Tone Discrimination of Cantonese and Mandarin between Tonal and Non-Tonal Speakers

As the world becomes more globalized, the number of interests aggregated in cross-linguistic studies. For the past 40 years, researchers had studied how linguistic background influences speech perception and it was often accepted that linguistic experience does play an important role. While numerous researchers had focused on how the perception of non-native languages was influenced by the listener's native language (L1) at segmental level, such as on areas regarding vowels and consonants, phonemes and word contexts to name but a few. It was only until recently that more attention were given at the suprasegmental level on the L1 influence to the listener's speech perception. Among the areas of suprasegmental level, it can be further differentiated depending on the language. Prosody, or suprasegmental features carry different functionality in different languages: for example, in Spanish, stress accents were used to differentiate between words, while in English, intonation of a word or throughout a sentence gives prosodic cues to listeners. On the other hand, for tonal languages, which are mainly found in East Asia and Africa, Pike (1976) stated that in for tonal languages, each syllable has "lexically significant, contrastive, but relative pitches". In this study, the interest lies in investigating between the perceptions of tonal languages, Cantonese and Mandarin, between both tonal (Cantonese and Mandarin) and non-tonal (English) language speakers.

There were previous researches studying cross-language perceptions between tonal and non-tonal language speakers. For instance, studies had been done to examine the relationship between the listener's own native language, whether it be tonal or non-tonal, and how it affected the perception of tones from a non-native tonal language.

However, conflicting results had been reported. Some researchers (e.g. Lee, Vakoch & Wurm, 1996; Wayland & Guion, 2004; van Dommelen & Husby, 2009) reported that listeners with more linguistic experience, that is more tonal contrasts in their L1, performed better in general than those whose native languages have no or very few tonal contrasts in perceiving non-native tones. For instance, Lee *et al.* (1996) concluded that Cantonese and Mandarin listeners performed better than English listeners when asked to differentiate non-native tones. Another example was Wayland & Guion (2004)'s study, which found that Thai-naïve Chinese listeners are better at discriminating Thai level tones than naïve English speakers.

On the other hand, there were also studies showing that linguistic experience do not necessarily facilitate non-native language perception, as shown by researches by Francis, Ciocca, Ma and Fenn (2008), Hao (2011), So (2005), So and Best (2010) etc. Rather than facilitation, these researchers propose that the influence of the listener's L1 may go both ways, i.e. it may aid or hinder their perception of foreign languages. To support this view, So and Best (2005) conducted a study on the perception of a tonal language (Mandarin) by Cantonese, Japanese and English listeners. Although both Cantonese and Japanese listeners had tonal language experiences: in the Cantonese inventory, there are more tone variations in terms of number (six tones) and acoustic dimension changes, while Japanese uses a pitch-accent system consisting of simple pitch shapes, the Cantonese listeners did not do better than the Japanese, thus reflecting that linguistic experience does not necessarily aid non-native language perception.

However, overall perception performances of non-native listeners in some researches may not be accurate on reflecting the actual influence of linguistic backgrounds, such as Lee *et al.*'s study in 1996, as they based the findings only on

overall performances. Thus, recent researches had started to analyze the specific patterns on how suprasegmental features from a particular L1 shapes the perception of the non-native tones (e.g. Francis *et al.*, 2008; Leung, 2008; So, 2005; So & Best, 2010). For example, Francis *et al.* (2008) conducted an experiment on the perceptual learning of Cantonese lexical tones by both native speakers of tonal (Mandarin) and non-tonal (English) language speakers. It was reported in the study that it was not adequate to define cross-language perception of lexical tones only based on lexical tone contrasts in the native languages, as both non-native listeners' perception performance on tone identification did not differ significantly. The researchers noted that the error patterns were quite different between the non-native groups. Also, going back to the study by So and Best (2008) on tone perception by Cantonese, Japanese and English listeners, the error patterns across all the three groups were different, where Cantonese listeners had less difficulty with Mandarin level and rising tones, of which the researchers explained as positive influence by Cantonese, and concluded that specific native languages affects perception of foreign languages specifically as reflected by the inconsistent error patterns across the experiment groups.

Thus, this study has 3 main objectives: 1) to investigate whether linguistic experience affects the perception of non-native tones; 2) to determine which tone pairs in Cantonese and Mandarin cause the most perceptual difficulty for non-native listeners (both tonal and non-tonal); and 3) to examine how different linguistic experiences with specific L1 affect the perception of non-native tones. More focus will be given on comparing the patterns of responses in relation to the listeners' native language, rather than focusing solely on the overall performance.

Two separate experiments will be conducted: Experiment 1 will focus on how Cantonese tones are perceived by 3 groups of listeners: native Cantonese, Mandarin and English listeners, both Mandarin and English listeners have no experience in Cantonese. Experiment 2 will focus on how Mandarin tones are perceived by 3 groups of listeners as well, this time by native Mandarin listeners, Cantonese listeners and English listeners. Both Cantonese and English listeners have no experience in Mandarin. These two similar experiments are conducted for a comparison between the performances and hopefully, to strengthen the findings and allow more conclusive results.

Tone inventories of Cantonese, Mandarin and English

Both Cantonese and Mandarin are tone languages, and pitch changes are linked specifically to lexical meanings in tonal languages (Pike, 1976).

In Cantonese, which is a lexical tone language, there are six contrastive tones (Bauer & Benedict, 1997; Fok Chan, 1974) which could be described in terms of Chao's (1930) 5-point pitch scale: Tone 1 (T1) *high level* [55]; Tone 2 (T2) *high rising* [25]; Tone 3 (T3) *mid level* [33]; Tone 4 (T4) *low falling* [21]; Tone 5 (T5) *low rising* [23]; Tone 6 (T6) *low level* [22].

Mandarin is also a lexical tone language (Yip, 2002). There are 4 contrastive tones (Bauer & Benedict, 1997; Li & Thompson, 1989; Norman, 1988): Tone 1 (T1) *high level* [55]; Tone 2 (T2) *mid rising* [35]; Tone 3 (T3) *falling-rising / dipping* [214]; Tone 4 (T4) *high falling* [51]. Also, it should be noted that a Mandarin sandhi rule gives a fifth contour (Li & Thompson, 1989), which is applied when a syllable with a *dipping tone* [214] (T3) is immediately followed by another T3 syllable, and the first T3 syllable will be produced as a *low falling* tone [21] instead.

Although both Cantonese and Mandarin are tonal languages, phonological similarities and differences could be found in several aspects. According to Leung (2008), Cantonese and Mandarin are distinguished from each other in terms of 4 aspects: 1) number and phonetic characteristics of tones; 2) perceptual correlates of tone; 3) types of error patterns; and 4) categoricity of tone perception, which means whether native speakers perceive tones categorically or continuously.

Table 1.

Tone contours and levels in Cantonese and Mandarin

Tone contour	Tone level	Cantonese		Mandarin	
Level	High	✓	[55]	✓	[55]
	Mid	✓	[33]	✗	-
	Low	✓	[22]	✗	-
Rising	High	✓	[25]	✓	[35]
	Low	✓	[23]	✗	-
Falling	High	✗	-	✓	[51]
	Low	✓	[21]	✗	-
Dipping (Falling-rising)	Low	✗	-	✓	[214]

Apart from the different number of tones, of which Cantonese has a richer tonal system in terms of numbers, for their phonetic characteristics, they differ in the number of contours and their respective magnitudes as seen in Table 1. For instance, both languages have level, rising and falling tones, while Mandarin has a dipping / falling-rising tone [214]. For level tones, all 3 level tones high [55], mid [33], low [22] are present in Cantonese, whereas only high-level [55] tone exists in Mandarin.

Regarding rising tones, both Cantonese and Mandarin has high rising tones [25], while Cantonese also has a low rising tone [23]. For falling tones, the two languages only have one falling tone each: Cantonese only has a low falling tone [21], while Mandarin only has a high falling tone [51]. Moreover, Cantonese does not have a dipping tone, whereas Mandarin has one [214].

For English, it could be regarded as a "stress-accent language" (Beckman, 1986). It is a non-tone language as no specific pitches that are linked lexically at word level. Apart from F_0 differences, vowel quality and length differences, which are more salient and consistent, weighs heavier for stress or unstressed syllables (Beckman, 1986; Fox, 2000; Gussenhoven, 2004). At phrasal level, English intonation use different pitch contours to represent prosodic meanings, for example, adopting the ToBI system as suggested by Beckman & Hirschberg (1994) and Beckman, Hirschberg, and Shattuck-Hufnagel (2005), rising contours (H-H%) are usually associated with the prosody for 'yes-no' questions, and falling contours (L-L%) for declarative statements. However, it should be noted that the intonation contours and pragmatic functions are not specifically linked. Although both Ladd (1996) and Ward and Hirschberg (1985) supported that intonation patterns are linked with specific prosodic meanings, the researches only covered the fall-rise contour, which is commonly used in expressing uncertainty, but there are no further evidence supporting direct linkages of specific pragmatic functions with other contours.

Perception of tones by different L1 listeners and prediction of performances

Apart from the influences from L1 tone inventories, psychoacoustic feature-based perspectives were also used for cross-language perceptions. Researches (Gandour,

1983; Gandour & Harshman, 1978; Halle, Chang, & Best (2004)) supported that among the acoustic dimensions of tone, such as F_0, duration and intensity, F_0 served as the main perceptual cue for lexical tones, which consisted of both pitch heights and contours. In addition, it was reported that relative weighing of different dimensions could be observed across listeners of different native languages, which will be listed specifically below.

Cantonese listeners. As mentioned above, it was well established that pitch height and contours play an important part in Cantonese tone perception. Tse (1973, 1978) and Vance (1976, 1977) further stated that tone levels play a more important part than contours in Cantonese. Gandour's study (1983) on tonal perception by a number of tonal speakers (Cantonese, Mandarin, Taiwanese, Thai) and English speakers also proved this finding. He found that "Cantonese listeners attached relatively more importance to the 'height' dimension than either Mandarin or Taiwanese listeners", although he noted that Cantonese listeners also pay attention to pitch contours.

Therefore, it could be predicted that Cantonese speakers would perform better at identifying tones with different contours, and also for tones pairs that differ in pitch heights but with the same contour patterns or directions than both Mandarin and English speakers (e.g. Cantonese level tones T1 & T3, T1 & T6, rising tones T2 & T5). As for perceiving Mandarin tones, native Cantonese listeners may have problems in differentiating tone pairs with T3, as there are no dipping tones in Cantonese.

Mandarin listeners. Similar to Cantonese, it was identified by researchers (Gandour, 1984; Massaro, Cohen, & Tseng, 1985; Moore & Jongman, 1997) that native Mandarin speakers adhere mainly to pitch heights and contours for tone identification. In addition, it was proved that native Mandarin speakers place much less importance on duration and intensity (Gandour, 1994; Tseng 1990), although it had been suggested by Ho (1976) and Howie (1976) that there are 'intrinsic durational differences' between Mandarin tones (e.g. T3, T2 > T1 > T4 in terms of duration). There is one particular cue that is essential in Mandarin tone perception to differentiate T2 and T3, which is the F_0 turning point (Moore & Jongman, 1997).

As Mandarin speakers have less experience with tone heights, it could be predicted that they perform worse on Cantonese tone pairs that differ mainly in tone heights (e.g. level tones T1 & T3, T1 & T6, T3 & T6 and contour tones T2 & T5). Also, as there are fewer variations of tones in Mandarin compared to Cantonese (e.g. magnitude of contour patterns), Mandarin speakers may have difficulty in discriminating tones that are not found in their native tone inventory (e.g. T3, T4, T5, T6).

English listeners. There are no lexical tones in English, and unlike Cantonese and Mandarin, intonation in English functions at the sentence level rather than the word level. Therefore, it is predicted that the English listeners will perform worse than the tonal speakers.

Summary

The present study aims to compare the perception of L2 tones by native speakers of both tone (Cantonese and Mandarin Chinese) and non-tone (English) languages.

As mentioned above, two experiments will be carried out to compare the specific L1 influence. In addition, more focus will be given on the error pattern analysis rather than the overall performance.

It was predicted that for Experiment 1, native Cantonese listeners would outperform Mandarin listeners, who would in turn perform better than English listeners. Results from Experiment 2 were expected to be similar, where native Mandarin listeners would be better than those speaking Cantonese, who were better than English listeners. These two assumptions were made from the hypothesis that L1 linguistic experience does play a role in non-native language perception. In addition, specific error patterns made by different L1 groups would be analyzed for a deeper insight on cross-language tone perception.

Method

Participants

Four groups of participants were recruited to participate in this experiment: 20 native speakers of Mandarin (8 M, 12 F), 20 natives speakers of English (11 M, 9 F), 20 native speakers of Cantonese (Mandarin naïve) (7 M, 13 F) and 20 native speakers of Cantonese (knows Mandarin) (8 M, 12 F). None of the Mandarin speakers were familiar with Cantonese or other tonal languages (by self-report through questionnaire). None of the English speakers were familiar with both Cantonese and Mandarin and other tonal languages (by self-report through questionnaire). The participants reported no speech or hearing impairments.

Stimuli

Experiment 1: Cantonese. Two syllables, /jau/ (/y:ou/ in IPA) and /se/ (/sɛ:/ in

IPA), each carrying six Cantonese tones, were used as test stimuli. One female native speaker of Hong Kong Cantonese was recorded reading the target syllables carrying six tones in a carrier phrase 我讀___字. "I read the word___.". A total of twelve tokens (2 syllables × 6 tones) were used (See Appendix A).

Experiment 2: Mandarin. Two syllables, /xi/ (/si/ in IPA) and /you/ (/yoʊ/ in IPA), each carrying four Mandarin tones, were used as test stimuli. One female native speaker of Mandarin (Beijing) was recorded reading the target syllables carrying four tones in a carrier phrase 我讀___字. "I read the word___.". A total of eight tokens (2 syllables × 4 tones) were used (See Appendix B).

Procedure

Same-different (AX) discrimination tasks were conducted, where stimuli, in this case tone pairs, are paired that they are either the same or different, and subjects were asked to discriminate if there are differences. All possible pairings of the tones and the two syllables (Experiment 1 - Cantonese: 6AA and 15AB pairs of the monosyllables; Experiment 2 – Mandarin: 4AA and 6AB pairs of the monosyllables) were randomized and presented to the participants, and the presentation order was counter-balanced in the AB pairs. For Experiment 1 (Cantonese), there were altogether 72 tokens (15 AB pairs × 2 syllables × 2 orders + 6 AA pairs × 2 syllables) with 60 AB pairs and 12 AA pairs, while for Experiment 2 (Mandarin), there are in total 32 tokens (6 AB pairs x 2 syllables x 2 orders + 4AA pairs x 2 syllables) with 12 AB pairs and 8 AA pairs. The stimuli were presented to the participants through a stereo headphone with the volume adjusted to a comfortable level in a quiet room. The participants were told that they would hear pairs of sounds from a certain language. After hearing a stimuli pair, the participants were asked to indicate whether the stimuli were the same or different on a charting form (see Appendix E and F).

The presentation of the stimuli was self-paced. However, the participants were only allowed to hear all the stimuli once. Before the tasks for both experiments, the participants were asked to listen to four examples chosen from the stimuli as demonstration of the task. Feedbacks were given only for the demonstration tasks. The whole experiment lasted for approximately 30 minutes.

Results (Experiment 1 – Cantonese)

The performance accuracy was calculated into percentages. Performances for the Cantonese tone pairs were shown in Figure 1. A two-way ANOVA was conducted to demonstrate the statistically significant differences between the 3 groups (Cantonese (Control), Mandarin and English) and the tone pair accuracies: the results revealed the main effects of Groups [$F(2,855) = 129.47$, $p < .001$] and of Tone Pairs [$F(14,855) = 35.79$, $p < .001$]. The interaction effect was also significant [$F(28,855) = 4.26$, $p < .001$]. Post-hoc comparisons (Bonferroni) found that there were significant differences between the 3 groups ($p < .001$): significant differences were found between the Cantonese and Mandarin group ($p < .001$), Cantonese and English group ($p < .001$), and also between the Mandarin and English group ($p < .001$).

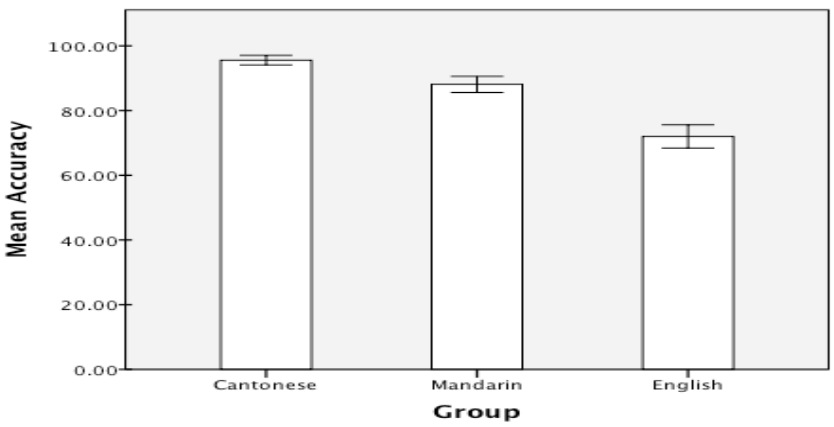

Figure 1. Accuracy percentage by different L1 groups

The results demonstrated that Mandarin listeners performed better than English speakers for overall performance, which may be attributed to the different L1 experience between the Mandarin and English listeners. This will be further addressed later in discussion.

Apart from the main effects, the performances for individual tone pair discrimination of each L1 group were examined in detail as shown by Figure 2 below.

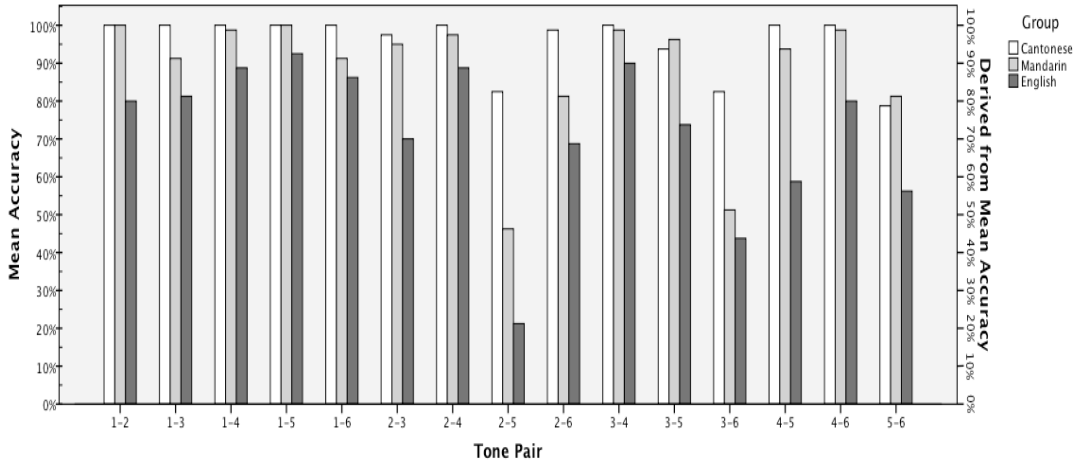

Figure 2. Error percentage of the 15 tone pairs in three L1 groups

It was shown in Figure 2 that the Mandarin listeners had particular difficulty with the level tone pair T3-T6 and also for pair T2-T5, while English listeners had most difficulty with the pairs T2-T5, T3-T4, T4-T5, and T5-T6, which mostly include contour tones.

Discussion (Experiment 1 – Cantonese)

Overall performance

Results showed that native Cantonese listeners (95.58%) performed better than Mandarin listeners (88.08%), who in turn also performed better than English speakers (72.0%) for the overall performance. The observation that native speakers of tone languages outperformed native speakers of a non-tonal language was consistent with

previous researches (e.g. Lee *et al.*, 1996, Wayland & Guion, 2004; van Dommelen & Husby, 2009), and for a more direct comparison, the result from Lee *et al.* (1996) also showed a decreasing accuracy trend for Cantonese listeners, followed by Mandarin and then English listeners. However, this observation conflicts with the study by Francis *et al.* (2008), where the performance of the English and Mandarin listeners on Cantonese tones perception did not differ significantly. The contrast between the Francis *et al.* (2008) study and the current study may probably be due to the differences in both stimuli and task, as their study used identification task, whereas the current study used same-different discrimination tasks. Francis *et al.* (2008) claimed that as the English and Mandarin listeners' performances for the identification of tones were comparable, thus they concluded that lexical tone contrasts between native languages was not the sole factor for cross-language perception of lexical tones.

Therefore, based on the findings for Experiment 1, the tentative conclusion is that the presence of lexical tone contrasts in native language, in this case Mandarin listeners, poses a positive influence on non-native tone discrimination. It would be further examined in Experiment 2 whether this conclusion could be generalized to other tonal languages, and whether it could be concluded that if tonal contrast experience in native languages plays a definitive role in cross language perception of tones.

However, it should be noted that the overall performances are not conclusive enough, thus error patterns from respective listener groups should be highlighted, as it is still unclear how the listeners' native language specifically affects their perception on non-native languages. This will be further discussed in the next session.

Cantonese tone perception by tonal language listeners. Referring to Figure 2 above (in Results), we can look at the performances for individual tone pairs for each group. The performance accuracy for tone pairs with T1 [55] was higher in general across all 3 groups, which is because it is far-separated from the other tones in acoustic space, i.e. T1 [55] has a high level pitch, while other tones have lower pitches and may have overlapping pitch values. As for the other tones, the error patterns and accuracy vary between each group, and only significant tone pair errors will be discussed below for conciseness.

It was shown that the native Cantonese listeners had the most difficulty with Cantonese tone pair T2-T5 (82.5%), T3-T6 (82.5%) and T5-T6 (78.75%). The particular difficulty with T2 and T5 was consistent with previous findings (e.g. Bauer & Benedict, 1997; Kei, Smith, So, Lau & Kapel, 2002; Fok, 1974), which reported that T2 [25] and T5 [23], both being rising tones, were easily confused. Also, the native speakers may find T3 and T6 difficult due to the very similar psychoacoustic properties, which was consistent to the findings by Li (2004), seeing T3 [33] and T6 [22] are both level tones and have similar average F_0s. As for the difficulty with the pair T5 and T6, although their contour direction is different, with T5 [23] being a low-rising tone and T6 [22] being a low-level tone, the contour difference is minimal, with a very similar average F_0.

For Mandarin listeners, who perceive Cantonese as a non-native tonal language, also had difficulty with the tone pairs T2-T5 (46.25%), T3-T6 (51.25%), and T5-T6 (81.25%) as native Cantonese speakers did, but with lower accuracy in general. However, interestingly, Mandarin listeners' performance at discriminating the tone pair T5 and T6 was comparable to the native Cantonese listeners (78.57%), despite of the advantages Cantonese listeners may have, such as having a richer linguistic

experience due to a larger tone inventory, or presuming that Cantonese speakers may be be able to adhere Cantonese tones to a lexicon, thus enabling better discrimination. Thus, this result maybe explained that the Mandarin listeners have also relied on the acoustic features of the tones. As found by Gandour (1983) and Francis *et al.* (2008), both Cantonese and Mandarin speakers attach importance to both acoustic dimensions of pitch height and contour when listening to a tonal language, which may explain why their performances were comparable. However, as it is only one example and with no further evidence, this could only be a very mild postulation.

Cantonese tone perception by non-tonal language listeners. As for English listeners, the performances for individual tone pairs were generally lower than that of tonal speakers. Apart from also having particular trouble with discriminating the tone pairs T2-T5 (21.25%), T3-T6 (43.75%), T5-T6 (56.25%) as the tonal speakers did, but with substantially lower accuracy, English listeners also tended to misidentify tone pairs T2-T3 (70.0%), T2-T6 (68.75%), T3-T5 (73.75%) and T4-T5 (58.75%). Comparing all the erroneous tone pairs for English listeners as seen in Table 2, it can be observed that the tone pairs all have similar tone heights and no specific pattern of adherence to direction of contour (e.g. rising vs falling vs level).

Table 2.

Erroneous Cantonese tone pairs perceived by English listeners and their corresponding pitch values

Erroneous tone pairs	Corresponding pitch values
T2-T3	[25] vs [33]
T2-T5	[25] vs [23]
T2-T6	[25] vs [22]

T3-T5	[33] vs [25]
T3-T6	[33] vs [22]
T4-T5	[21] vs [23]
T5-T6	[23] vs [22]

Explaining these error patterns in terms of acoustic features, this finding is consistent to that of Gandour (1983) and Gandour and Harshman (1978), who also found that English speakers tended to focus more on pitch height than contours. Furthermore, although there are pitch contours, or intonations in English, which some researchers (e.g. So, 2005; So & Best, 2008) postulate may be comparable to tones, they usually function at sentence level for prosody (e.g. implications for statement or questions), rather than at the word-level. Thus, English speakers may not be able to utilize their experience with intonation contours to the perception of lexical tones.

Summary for Experiment 1

Comparing the patterns across the three different language groups, it could be seen that there was a slightly similar pattern of errors for certain tone pairs (T2-T5, T3-T6, T5-T6) with different accuracies (Cantonese > Mandarin > English). The tone pair accuracies for both Cantonese and Mandarin listeners were mostly comparable except for a few (T2-T5, T2-T6, T3-T6), while English listeners had more difficulty in tone discrimination which was reflected in terms of number of erroneous tone pairs and considerably lower accuracy.

Relating to the three objectives for the current study, temporary conclusions could be made: First, on whether linguistic experiences affect non-native tone perception, it could only be concluded that L1 experiences does play a role in cross-language

perception, but it cannot be determined whether the influence is positive or negative, and to what extent the importance of linguistic background is. For the second objective, for difficulty in perception for Cantonese tones, T2-T5 and T3-T6 was found to be most perceptually difficult for both tonal and non-tonal listeners. As to the question of how specific L1 experiences influence non-native tonal perception, for Mandarin, it was expected that Mandarin listeners would be able to distinguish Cantonese tones that are similar to those in the Mandarin inventory relatively well (e.g. High-level T1 [55], High-Rising T2 [25], Low-Falling T4 [21]), and worse for those tones that are dissimilar (e.g. Mid-Level T3 [33], Low-Rising T5 [23], Low-Level T6 [22], as there are no corresponding pitch heights in Mandarin tones) (refer to Table 1 in Introduction). Yet, results show that Mandarin listeners did not perform significantly poorer for all tones that are 'dissimilar' to those in Mandarin inventory, but rather only for certain tone pairs. In addition, it was hypothesized that since English listeners has no experience with lexical tones, their performance for all tone pairs would be worse in general, but as it turns out, although the performance was lower across all tone pairs, the English listeners found some tone pairs significantly more difficult to distinguish rather than all tone pairs.

Therefore, this may indicate that rather relying merely on linguistic experience, non-native listeners also adhere to other aspects for cross-language perception, to which there are two adopted theories from recent studies: some researchers advocated 'category assimilation' (e.g. the Perceptual Assimilation Model (PAM) proposed by Best in 1995, p.194), that is non-native listeners "have a strong tendency to perceptually assimilate non-native speech categories to the native categories that they perceive as the most similar". On the other hand, recent researches (e.g. Wang, Behne, Jongman & Sereno, 2004) suggested another possible approach: the study

investigated the hemispheric lateralization of Mandarin tones by tonal and non-tonal speakers, and suggested that acoustic features may play a more pivotal role in cross-language perception rather than existence of the native tone itself, that is the specific acoustic features from a native language affects the acoustic features used when listening to a foreign language.

The study will continue to analyze the results from Experiment 2 (Mandarin as stimuli). Comparisons were be made between Experiment 1 and 2 for further discussion and conclusions.

Results (Experiment 2 – Mandarin)

Figure 3 shows the performance accuracy percentages for Mandarin tone pairs. A two-way ANOVA was conducted to demonstrate the statistically significant differences between the 3 groups (Mandarin (Control), Cantonese and English) and the tone pair accuracies: the results revealed the main effects of Groups [$F(2,342) = 65.66$, $p < .001$] and of Tone Pairs [$F(5,342) = 38.28$, $p < .001$]. The interaction effect was also significant [$F(10,855) = 9.84$, $p < .001$]. Post-hoc comparisons (Bonferroni) found that there were significant differences between the 3 groups ($p < .001$): significant differences were found between the Mandarin and Cantonese group ($p < .001$), Mandarin and English group ($p < .001$), and also between the Cantonese and English group ($p < .001$).

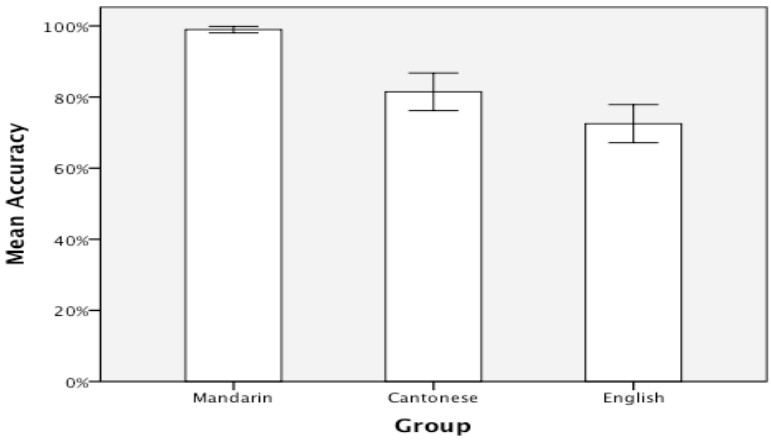

Figure 3. Performance accuracy percentage by different L1 groups

Apart from the main effects, the performances for individual tone pair discrimination of each L1 group were examined in detail as shown by Figure 4 below.

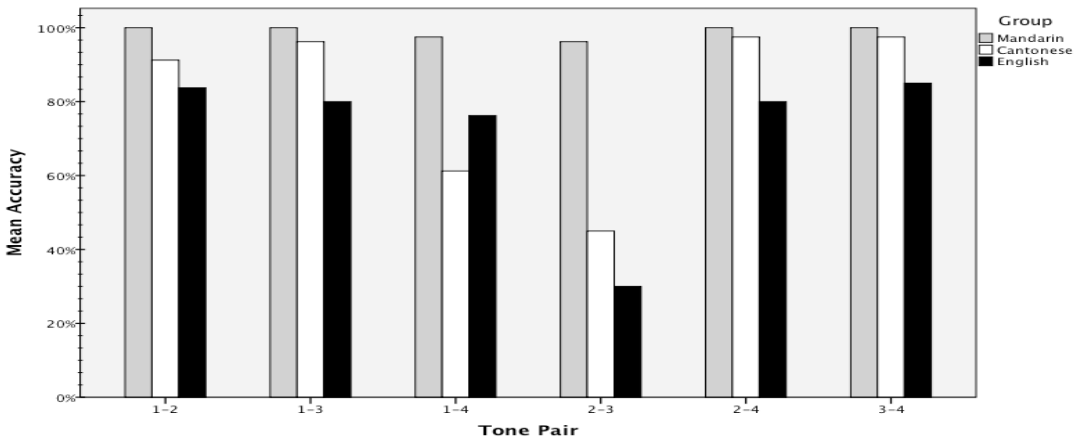

Figure 4. Error percentage of the 6 tone pairs in three L1 groups

To conclude, the results demonstrated that while Mandarin listeners performed the best out of the three groups, Cantonese listeners performed better than English speakers for overall performance, which can be attributed again, similarly to Experiment 1, to the different L1 experience between the Cantonese and English listeners. As for the individual tone pair discrimination, it was shown in Figure 4 that the Cantonese listeners had particular difficulty with the tone pair T1-T4 and T2-T3, while English listeners had difficulty with most tone pairs, but also with T2-T3 especially.

Discussion (Experiment 2 – Mandarin)

Overall Performance

Overall performance results showed that native Mandarin listeners (98.95%) performed better than Cantonese listeners (81.46%), while Cantonese listeners performed better than English speakers (72.50%). Again, native speakers of tone languages outperformed native speakers of a non-tonal language, which was consistent with the result from Experiment 1 and previous researches (e.g. Lee *et al.*, 1996; So & Best, 2008). Apart from using Cantonese as stimuli as mentioned previously in Experiment 1, Lee *et al.* (1996) also conducted an experiment with Mandarin as stimuli that showed a decreasing accuracy trend between the native Mandarin groups, followed by Cantonese listeners and then English listeners (Native tonal listener > Non-native tonal listener > Non-native non-tonal listener). In addition, similar results were found in So and Best (2008)'s study, where Cantonese listeners outperformed English speakers in identifying Mandarin tones.

Mandarin tone perception by tonal language listeners. Referring to Figure 4 (found in Results), native Mandarin speakers had no or very little difficulty with discriminating the tone pairs. As for Mandarin naïve Cantonese listeners, the tone discrimination performances were comparable to that of Mandarin listeners except for Mandarin tone pairs T1-T4 (61.25%) and T2-T3 (45.0%). This finding is consistent with most researches (e.g. Hao, 2011; So, 2005; So & Best, 2008). For the tone pair T1 [55] and T4 [51], Cantonese listeners may not have difficulty with T1 [55], since the Cantonese T1 has the exact same pitch value (both height and contour) (refer to Table 1 in Introduction). Instead, confusion may occur for T4. Although in Cantonese there is also a Low-Falling tone [21], the falling height is much smaller

than High-Falling Mandarin T4 [51]. Presumably, Cantonese listeners could adhere the Mandarin T4 [51] to Cantonese T4 [21] as they both have the same direction of contour (falling); however, it was perceived as a level tone T1 [55] instead, so the Cantonese listener may have adhered more to the height dimension, as both T1 and T4 has the same pitch onset. This view is supported by Hao (2011), that Cantonese speakers may have assimilated both Mandarin T1 and T4 to Cantonese T4.

For the other erroneous Mandarin tone pair T2-T3, there was one similar Cantonese tone correlate T2 [25] for Mandarin T2 [35], whereas there were no correlates in the Cantonese tone inventory for the Mandarin Dipping-Tone T3 [214]. It could be explained that since there are no dipping tones in Cantonese, Cantonese listeners might extract the rising part of T3 [14] and then perceived Mandarin T3 as a rising tone, such as T2 [35]. Moreover, both T2 [35] and T3 [214] have similar pitch onset and offset values. Therefore, due to the linguistic background (or lack of) and similar acoustic features (both height and contour dimensions), Cantonese listeners may easily confuse between T2-T3.

Mandarin tone perception by non-tonal language listeners. For English listeners, their performance was again generally lower compared to that of the tonal listeners, and they also experienced particular difficulty with the Mandarin T2-T3 (30.0%), which was expected as a lot of studies suggested that both T2-T3 were commonly found to be difficult for non-tonal language learners (e.g. English, Australian and Dutch) of Mandarin according to So and Best (2008).

One tone pair to note is the Mandarin T1-T4. It was shown above that Cantonese listeners had particular difficulty with both T2-T3 and T1-T4. However, the same

situation was not observed for English listeners. The English listeners' performances for the tone pair T1-T4 (76.25%) was comparable to that of other tone pairs (mean accuracy for English listeners: 72.50%), which is consistent to the study from Hao (2011) and Qin and Mok (2011). The English listeners rarely had confusion for Mandarin T1 and T4 tones. As mentioned before, English has no lexical tone system and the prosodic elements only function at sentence level instead of word level, thus the possibility of English listeners adhering to the intonation contours (e.g. falling contour (L-L%) for declarative statements to Mandarin T4 [51]) should be presumably low. Some studies (e.g. So, 2005; So & Best, 2008) tried to map intonation to lexical tones, but as there is no evidence that English intonation could be assimilated to native language categories, this view will not be implemented in this study. Therefore, linguistic background may not affect English listeners on the perception of tones, but rather, relied more on acoustic feature perception, as indicated by the better performance in T1-T4.

Summary for Experiment 2

Both Cantonese and English listeners had particular difficulty with the tone pairs T2-T3, but only Cantonese speakers had more difficulty in differentiating T1-T4. This finding proved that having tonal language experience does not necessarily facilitate non-native language perception, or even hinder cross-language perception, as suggested by Best (1995), such as the possible categorical assimilation of both Mandarin T1 and T4 to Cantonese T4 by Cantonese listeners. This also indicates that there are other factors, such as acoustic features (e.g. the acoustic similarity of T2 & T3) affecting such perception. This finding is consistent with both Hao (2011) and Qin and Mok (2011). In Hao's research, both Mandarin naïve Cantonese and English

speakers were asked to perceive and learn Mandarin and found that Cantonese speakers found the T1 & T4 pairs harder to perceive than English speakers. Furthermore, after learning, although Cantonese speakers had improved in differentiating T1 & T4, the improvement for T2 & T3 was not as significant as English speakers. Similar to the current study, Hao concluded that apart from the influence of L1 backgrounds, other factors such as phonological relationship and psychoacoustic features also play important roles in non-native language perception.

Conclusion

Comparing the results from both Experiment 1 and 2, the objective findings were summarized. First, regarding the effect of linguistic background, the L1 groups may seem to share similar overall performances. However, upon examination of individual tone pair performances, all the groups have different performances on specific tone pairs in terms of accuracy and error patterns. Although some of the erroneous tone pairs were shared among the groups, it was usually due to psychoacoustic similarities. For example, in the study, it was observed that Cantonese speakers also had minor difficulty with discriminating tone from their own native language, which indicated that apart from linguistic experience (e.g. adhering tone to lexical aspects), acoustic features also contributes to language perception. The other unshared tone pair errors between language groups reflected that linguistic background (i.e. native language) does have specific patterns of influence on perceiving non-native languages.

After establishing that one's native language does have specific influences on the perception of non-native languages, the next objective was to investigate the most perceptually difficult tones for each language group and how the listener's native

language specifically cause the difficulty: for Cantonese tones perception, it was found that Mandarin speakers found it harder to discriminate level tones, as Mandarin level tones do not have tonal contrasts, so they are more sensitive to pitch direction than pitch heights. On the other hand, for Mandarin tones perception, Cantonese speakers did not have particular difficulty with a type of tone, but it was observed that they adhered more to pitch heights than direction, which is because in the Cantonese inventory, most tones share the same direction and thus, pitch height is highly important for tone discrimination. Whereas, for both tonal languages, English speakers were observed to place more importance on pitch heights than contours or directions, and with much lower accuracy as English has no tonal system and thus, they have difficulty utilizing the acoustic dimensions (e.g. pitch height and contour) for tone discrimination.

Further considerations

It has been concluded in the current study that both linguistic experience and specific acoustic features from the native language play roles in non-native language perception. However, it is still unclear of the weighing between various aspects, for example, if adherence to acoustic features plays a more pivotal role. Dimension weighing within acoustic features is another area to be investigated on. For instance, dissimilarity tasks could be carried out to examine the weighing of acoustic dimensions (e.g. Francis *et al.*, 2008; Gandour, 1983; Mok, 2011). Apart from acoustic factors, other aspects are currently postulated to be included in non-native language perception as well, such as categorical assimilations between the native and non-native language categories (e.g. Best's PAM model, 1995), or using neural studies of hemispheric lateralization to study language and/or acoustic involvement in tone perception (e.g. Klein, Zatorre, Milner & Zhao, 2000; Wang *et al.* 2004).

Acknowledgements

I would like to thank Dr. Lydia So from the Department of Speech and Hearing Sciences of The University of Hong Kong especially for the advice and supervision on this dissertation. I also wish to thank my schoolmates and friends for their generous help and support.

References

Bauer, R. S., & Benedict, P. K. (1997). Modern Cantonese phonology. Berlin: Mouton de Gruyter.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in cross-language research,* 171-204. New York.

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

Chao, Y. R. (1930). A system of tone letters. *La Maître phonétique*, *45*, 24–27.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non- tone language speakers. *Journal of Phonetics, 36* (2), 268-294.

Fok Chan, Y. Y. (1974). A perceptual study of tones in Cantonese. Hong Kong: Centre of Asian Studies, University of Hong Kong.

Gandour, J. T. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*(2), 149–175.

Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: a multidimen-sional scaling investigation. *Language and Speech, 21*(1), 1–33.

Hao, Y.C. (2011). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics, 40*, 269-279.

Ho, A. T. (1976). The acoustic variation of Mandarin tones. *Phonetica, 33*, 353-367.

Howie, J. M. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge: Cambridge University Press.

Klein, D., Zatorre, R. J., Milner, B. & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage*, *13*, 646 - 653.

Lee, Y., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*, 527-542.

Leung, A.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, P. Y. (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.

Kei, J., Smith, V., So, L. K. H, Lau, C. C., Kapel, K. (2002). Assessing the accuracy of production of Cantonese lexical tones: a comparison between perceptual judgement and an instrumental measure. *Asia Pacific Journal of Speech, Language and Hearing, 7*, 25-38.

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

Qin, Z. & Mok, P. (2011) Discrimination of Cantonese tones by Mandarin, English and French speakers. In *Proceedings of the Psycholinguistic Representation of Tone Conference (PLRT)*, 50-53. Hong Kong.

Qin, Z. & Mok, P. (2011) Perception of Cantonese tones by Mandarin, English and French speakers. In *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, 1654-1657. Hong Kong.

So, C. K. (2005). The influence of L1 prosodic background on the learning of Mandarin tones: patterns of tonal confusion by Cantonese and Japanese native

listeners. In: C. Gurski. (Ed.), *Proceedings of the 2005 CLA Annual Conference.*

So, C. K., & Best, C.T. (2010). Cross-language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences. *Language and Speech, 53*, 273-293

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics*, *25*, 449 - 466.

van Dommelen, W. A., & Husby, O. (2009). The perception of Norwegian word tones by Chinese and German listeners. *Journal of the Acoustic Society of America, 125*, 2773-2773.

Wayland, R. P., Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*, 681-712.

Yip, M. (2002). *Tone.* Cambridge: Cambridge University Press.

Appendix A

Cantonese Word Stimuli for Experiment 1

Stimuli 1 - /jau/

Tone	Phonetic Transcription	Word
1	/jau/1	優
2	/jau/2	柚
3	/jau/3	幼
4	/jau/4	由
5	/jau/5	友
6	/jau/6	又

Stimuli 2 - /se/

Tone	Phonetic Transcription	Word
1	/se/1	些
2	/se/2	寫
3	/se/3	卸
4	/se/4	蛇
5	/se/5	社
6	/se/6	射

Appendix B

Mandarin Word Stimuli for Experiment 2

Stimuli 1 - /xi/

Tone	Phonetic Transcription	Word
1	/xi/1	西
2	/xi/2	習
3	/xi/3	洗
4	/xi/4	細

Stimuli 2 - /you/

Tone	Phonetic Transcription	Word
1	/you/1	優
2	/you/2	由
3	/you/3	友
4	/you/4	又

Appendix C

Language Background Questionnaire (English)

Questionnaire on Language Background

1	What is your age?	
2	Where are you from? (Country of origin)	
3	Have you lived in another country(s) for more than a year? - If yes, please specify and see Q4 - If no, please skip to Q5	
4	What is/are the major language(s) spoken in that country?	
5	What is/are your native language(s)? (If you speak a certain dialect, please specify)	
6	Do you know other language(s)? - If yes, please specify and see Q7 & 8 - If no, please skip to Q9	
7	At what age did you start learning it/them?	
8	Please rate your language proficiency at conversation level: 1 = understand but cannot speak 2 = understand and can speak with great difficulty 3 = understand and speak but with some difficulty 4 = understand and speak comfortably, with little difficulty 5 = understand and speak fluently like a native speaker	

Appendix D

Language Background Questionnaire (Chinese)

關於語言背景問卷

1	請問你的年齡是?	
2	你在哪裡長大？	
3	你曾否在外地生活超過 1 年？ - 有 : 請列明及回答 題 4 - 沒有: 請到 題 5	
4	那國家的主要語言是？	
5	你的母語是？ (如過有方言的話，請列明)	
6	你懂其他語言嗎？（可多於一種） - 有 : 請列明及回答 題 7,8 - 沒有: 請到最後	
7 (接題 6)	你從幾多歲開始學習哪語言？	
8 (接題 6)	請說明你的語言程度 (交談程度，不包括讀寫) ： (基本 / 中階 / 進階 / 流利)	

Appendix E

Participant Charting Form (English)

Instructions:
You will hear some recordings in Cantonese and Mandarin. There will be two sentences in each recording.
Listen carefully to determine if the **third word/syllable** in both sentences **may or may not sound different from each other.**
Please only count the 3rd word and not others

Please indicate in the boxes below:
If you think they sound the same, put a ✔
If you think they sound different, put a ✖
Please listen to track Example. 1, 2, 3, 4 for demonstration.

Example 1	✔ (They sound the same.)	Example 2	✖ (They sound different.)
Example 3	✔ (They sound the same.)	Example 4	✖ (They sound different.)

Part 1: Cantonese

1		10		19		28	
2		11		20		29	
3		12		21		30	
4		13		22		31	
5		14		23		32	
6		15		24		33	
7		16		25		34	
8		17		26		35	
9		18		27		36	

Feel free to take a rest before starting again.

37		46		55		64	
38		47		56		65	
39		48		57		66	
40		49		58		67	
41		50		59		68	
42		51		60		69	
43		52		61		70	
44		53		62		71	
45		54		63		72	

Part 2: Mandarin

1		5		9		13	
2		6		10		14	
3		7		11		15	
4		8		12		16	

Feel free to take a rest before starting again.

17		21		25		29	
18		22		26		30	
19		23		27		31	
20		24		28		32	

Appendix F

Participant Charting Form (Chinese)

你將會聽到一些廣東話 (Cantonese Part 1) 和普通話 (Mandarin Part 2) 錄音，每個錄音裏有兩句句子。 請***只留意***兩句句子中的***第三個字***，他們聽上去可能會有分別。 **＊請只記錄第三個字有沒有分別。＊** 聆聽每段錄音後，請在相應的題號旁的方格裏填寫 ✔(代表一樣的字) 或 ✖ 填(代表不一樣的字)。
請先聆聽 Example. 1，2，3，4 作例子:

例子 1 (Cantonese e.g. 1)： ✔	例子 2 (Cantonese e.g. 2)：✖
例子 3 (Cantonese e.g. 3)： ✔	例子 4 (Cantonese e.g. 4)：✖

第一部份：**廣東話**

1		10		19		28	
2		11		20		29	
3		12		21		30	
4		13		22		31	
5		14		23		32	
6		15		24		33	
7		16		25		34	
8		17		26		35	
9		18		27		36	

請休息一會，準備好的時候再開始!

37		46		55		64	
38		47		56		65	
39		48		57		66	
40		49		58		67	
41		50		59		68	
42		51		60		69	
43		52		61		70	
44		53		62		71	
45		54		63		72	

第二部份：**普通話 (Mandarin Part 2)**

1		5		9		13	
2		6		10		14	
3		7		11		15	
4		8		12		16	

請休息一會，準備好的時後再開始!

17		21		25		29	
18		22		26		30	
19		23		27		31	
20		24		28		32	