TONAL ASSIMILATION PATTERNS OF CANTONESE L2 SPEAKERS OF MANDARIN IN THE PERCEPTION AND PRODUCTION OF MANDARIN TONES[*]

Anna S. Leung
University of Toronto

1.   Introduction

Cantonese and Mandarin are both tonal languages, in which variations in pitch are used to convey the semantics of lexical items. Cantonese, which has six contrastive lexical tones (Flynn 2003), has a larger tonal inventory than Mandarin, which has four contrastive lexical tones (Howie 1976; Blicher, Diehl, & Cohen 1990).   A comparison of the overall acoustic shapes between Cantonese and Mandarin tones shows notable similarities and differences. For example, both languages have high level and high rising tones, but Cantonese has additional level tones and an additional rising tone, while Mandarin has a high falling tone that does not have a precise Cantonese counterpart. The various tonal similarities and differences between these two tonal inventories suggest that the phonetic realization of Mandarin tones by native (L1) Cantonese-speaking second-language (L2) learners of Mandarin would be influenced by the Cantonese speakers' pre-existing tonal repertoire. As such, the acquisition of Mandarin tones by Cantonese speakers poses an interesting topic for L2 theory.

Previous studies on lexical tone perception have mainly focused on the perception of non-native tones by naïve listeners whose native language does not utilize tone contrastively (e.g., Hallé, Chang, & Best 2003), or by naïve listeners who have had prior experience with a tone language (e.g., Lee, Vakoch, & Wurm 1996). The current study departs from these perceptual studies in that it investigates the perceptual assimilation of non-native tones by listeners who not only have L1 tonal experience, but who have also received formal instruction on the target non-native language as L2 learners; namely, Cantonese learners of Mandarin. This study thus acts as a precursor for further study in the field of lexical tone perception and production by L2-learners of a target tonal language who have also had L1 tonal experience. By examining the perception and production of Mandarin tones by Cantonese learners of Mandarin, the results of the current study will advance our knowledge and understanding of L2 tone acquisition and complement existing literature on non-native tone perception.

2. Objectives

The objectives of the current study are threefold: (1) to determine which Mandarin tones cause the most perceptual and production difficulties for Cantonese learners of Mandarin; (2) to determine the most plausible Cantonese counterparts of Mandarin tones; and (3) to evaluate whether the assimilatory predictions proposed for the perception of segments in one of the prevailing models of non-native speech perception, namely Best's Perceptual Assimilation Model (1995), can be extended to explain the perceptual assimilation patterns of non-native lexical tones as perceived by Cantonese learners of Mandarin.

3. Linguistic Background

Figure 1 displays the normalized fundamental frequency contours (falling within the range of Chao's (1947) 5-point pitch scale) for the six tones of Cantonese (C1, C3, C6 are the high level, mid level, and low level tones, respectively; C2 is the high rising tone; C5 is the low rising tone) (shown with smooth lines), and for the four tones of Mandarin (M1 is the high level tone; M2 is the high rising tone; M3 is the low falling rising tone; and M4 is the high falling tone) (shown with dotted lines), respectively. Note that both tonal inventories are manifested on the syllable [ma].
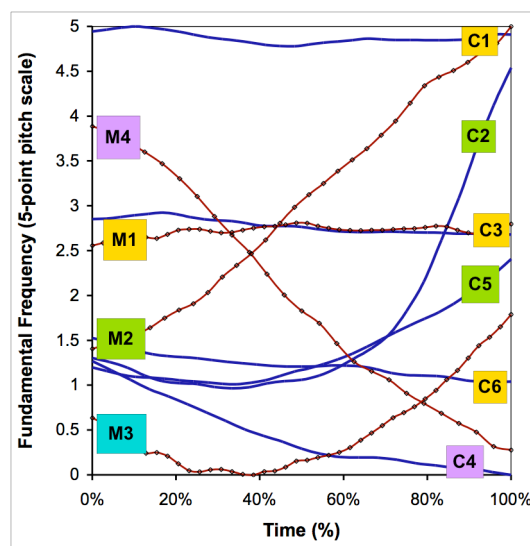

Figure 1. Cantonese and Mandarin tones

Cantonese and Mandarin are distinguished from each other in terms of: (i) number and phonetic characteristics of tones; (ii) perceptual correlates of tone; (iii) types of error patterns; and (iv) categoricity of tone perception (i.e., whether a set of tones are perceived categorically or continuously by native speakers).

Table 1. Number and nature of tones

TONAL SHAPE	TONES	CANTONESE	MANDARIN
■ Level	High Mid Low	✓ ✓ ✓	✓ ✗ ✗
■ Rising	High Low	✓ ✓	✓ ✗
■ Falling	High Low	✗ ✓	✓ ✗
■ Dipping	Low falling rising	✗	✓
	Total # of tones	= 6	= 4

As seen in Table 1, Cantonese has the richer tonal system, with six tones, versus Mandarin, which has four tones. In both cases, there are level tones, rising tones, and falling tones. However, the number of each of these tonal contours and the magnitude of these tonal contours are different in some cases. Referring to Table 1, each language has a high level tone, but the Cantonese tonal inventory has two additional level tones (mid level and low level), which do not exist in Mandarin. In terms of rising tones, both Cantonese and Mandarin each have a high rising tone, but Cantonese has an additional low rising tone, while Mandarin has a rising tone that has an initial dip, creating a low falling rising tone. Finally, with regards to falling tones, Cantonese has only one falling tone, the low falling tone, whereas Mandarin has a high falling tone.

Table 2. Perceptual correlates of tone

CORRELATE	CANTONESE	MANDARIN
F_0 height	✓	✗
F_0 shape / contour	✓	✓
Duration / Length	✗	✓
Amplitude / Intensity	✗	✓

Referring to Table 2, there are more relevant perceptual correlates of tone for Mandarin than for Cantonese. Native Cantonese speakers rely mainly on tonal height and shape for native tone identification (Fok 1974), whereas native Mandarin speakers rely mainly on tonal shape, tonal duration, and tonal amplitude for native tone identification (Howie 1976; Whalen & Xu 1992). Overall, in Mandarin, M4, the high falling tone, is the shortest tone, and M3, the low falling rising tone, is typically the longest (Whalen & Xu 1992).

Tonal confusions tend to arise when the acoustic properties of the tones are phonetically similar. For native Cantonese speakers, tonal confusions arise mainly with regards to the three level tones and the two rising tones since these two sets of tones differ in height and in magnitude of the contour, respectively (Fok 1974). For native Mandarin speakers, tonal confusions do not typically

arise with regards to tone height, but the acoustic similarities between the Mandarin high rising tone and the Mandarin low falling rising tone are sufficient to cause identification errors between these two tones (Whalen & Xu 1992).

Finally, Mandarin tones are perceived categorically because the shapes of the tones are more distinct from each other. Additional perceptual cues related to Mandarin tone perception, such as duration and intensity, which do not correlate with Cantonese tone perception, may further facilitate the identification of Mandarin tones by making the contrasts between the tones in the Mandarin inventory more salient and robust. Cantonese tones, on the other hand, are mainly perceived continuously because several Cantonese tones share very similar acoustic shapes, differing only along one dimension (e.g., tone height or magnitude of tone rise), and have fewer perceptual correlates with which to rely upon for accurate tone identification.

4.   Hypotheses

In general, the mechanism by which native Cantonese speakers perceive and produce Mandarin tones is expected to be a function of the degree of acoustic contrast between the target Mandarin tone and the tones which exist in the Cantonese tonal space. Specifically, if perceptual assimilations do occur, then by extending the patterns of perceptual assimilation posited for segments in Best's (1995) Perceptual Assimilation Model (PAM) to the assimilation of non-native tones, the prediction is that the Mandarin tones, as perceived by native Cantonese speakers, will be assimilated to native tonal categories rather than to an uncategorizable speech sound or to a non-speech sound. The latter two possibilities for tonal assimilation are ruled out in this case because tones serve the same linguistic function in Cantonese as they do in Mandarin, namely to distinguish lexical meanings. Consequently, Cantonese learners of Mandarin are not expected to perceive the non-native Mandarin tones as uncategorizable speech sounds; nor are they expected to assimilate the Mandarin tones as a non-speech sound (e.g., a bird chirp). Thus, under the assumption that the non-native Mandarin tones will be mapped onto existing Cantonese tones by native Cantonese speakers, the non-native Mandarin tones can be perceived in one of three ways (based on Best 1995: 194): (i) as a good exemplar of a native Cantonese tone; (ii) as an acceptable, but non-ideal exemplar of a Cantonese tone; or (iii) as a notably deviant exemplar of a Cantonese tone.

Based on the heights and shapes of the four Mandarin tones in Figure 1, the closest Cantonese tonal counterpart of M1, the Mandarin high level tone, appears to be C3, the Cantonese mid level tone, which has the most similar height and shape as M1. Although both the high level (C1) and low level (C6) Cantonese tones share the same level shape as the Mandarin high level tone (M1), based on Figure 1, M1 is still considered a better exemplar of the native Cantonese mid level tone, C3, than the native Cantonese high and low level tones (C1 and C6, respectively) because M1 and C3 also share a comparable height with each other. The high level Mandarin tone, however, is still considered to be an acceptable, although non-ideal, exemplar of the high level

Cantonese tone and the low level Cantonese tone, based on similarities in tone shape.

The Cantonese tonal counterpart of M2, the Mandarin high rising tone, appears to be C2, the Cantonese high rising tone. Although the high rising tone in Mandarin has a similar contour as the low rising tone in Cantonese (C5), the magnitude of the slope associated with M2 is noticeably greater than that of C5, and as such, the slope of M2 is more analogous to the slope associated with C2. Thus, based on the similar magnitudes and shapes of slope shared by M2 and C2, M2 is considered to be a better exemplar of C2, although M2 is still considered to be an acceptable exemplar of C5.

Meanwhile, M3, the Mandarin low falling rising tone, does not seem to be a good exemplar of any specific Cantonese tone because Cantonese does not have a corresponding dipping tone. However, Cantonese does have falling and rising tones, which are both components of the dipping tone. Referring to Figure 1, the falling portion of the Mandarin dipping tone (M3) can be considered to be an acceptable exemplar of several different Cantonese tones when only the onset portion of the Cantonese tone is examined. Specifically, the falling portion of M3 seems to descend to a similar degree as the initial descending slope of C2, C5, and C4, that is, the Cantonese high rising, low rising, and low falling tones, respectively. Arguably, the falling component of the Mandarin dipping tone may also be considered to be an acceptable exemplar of the slight initial falling slope of C6, the low level tone. However, if a scale were to be devised to rank the levels of acceptability between C2, C5, C4, and C6, then C6 is predicted to be less acceptable as an exemplar because of its overall level shape. If only the rising portion of the Mandarin dipping tone (M3) is under consideration, then M3 would be an acceptable exemplar of either of the two Cantonese rising tones, C2 or C5. By integrating the comparative analyses corresponding to the two elements of the dipping tone: the initial falling portion and the eventual rising portion, M3 is considered to be a more acceptable exemplar of both C2 and C5 than either of C4 or C6. Unlike the two rising tones in Cantonese (C2 and C5), both the low falling tone (C4) and the low level tone (C6) do not consist of a final rising element which is associated with the dipping tone in Mandarin. Moreover, although M3 seems to be in closer proximity to C4 than the other potential Cantonese counterparts in terms of height, (at least during the first half of its tonal contour), M3 still appears to correspond more closely to the two rising tones in Cantonese in terms of overall contour shape. Of these two rising tones, M3 would most likely be a better exemplar of the Cantonese high rising tone (C2), even though M3 is in closer proximity to the Cantonese low rising tone (C5). The interpretation that M3 is a better exemplar of C2 is based on the idea that if native Mandarin speakers tend to perceive and misidentify M3 as M2 (despite their apparent pitch differences in Figure 1), then Cantonese speakers who are acquiring Mandarin tones non-natively, are also likely to misperceive M3 as M2. If such misperception occurs, then the Cantonese speakers will likely perceive the dipping tone as the Mandarin high rising tone, which is a better exemplar of the Cantonese high rising tone than the Cantonese low rising tone. Consequently, based on this

rationale, the more acceptable Cantonese tonal counterpart of M3 is considered to be C2 rather than C5, although M3 is still considered to be a reasonable exemplar of C5.

Finally, M4, the Mandarin high falling tone, appears to be a notably deviant exemplar of the available tones in Cantonese. Referring to Figure 1, a clear counterpart to the Mandarin high falling tone does not exist within the Cantonese tonal repertoire in terms of slope. Although a low falling tone does exist in Cantonese (C4), its slope does not fall as abruptly, nor as sharply as the Mandarin high falling tone to be deemed a convincing tonal counterpart of the Mandarin high falling tone. Furthermore, the initial height of the Cantonese falling tone is noticeably lower than that of the Mandarin falling tone: C4 starts at approximately 1.25 on the 5-point pitch scale, whereas M4 starts much higher at approximately 3.9. Because M4 does not have a direct Cantonese counterpart in terms of slope, the abrupt falling slope that characterizes M4 may actually be less salient to native Cantonese speakers and, therefore, more difficult to perceive. M4 may, however, be an acceptable, though non-ideal, exemplar of the high level tone in Cantonese (C1) based on their comparable heights: the relative frequency of M4 is approximately 3.9 and the relative frequency of C1 is 5. By mainly focusing on the initial height of M4 rather than on its sharp and abruptly falling slope, which is absent in the Cantonese tonal inventory, native Cantonese speakers may perceive the high falling tone simply as a high level tone. In this case, M4 would most likely be perceptually assimilated to the high level Cantonese tone based solely on the height dimension. Hence, although M4 would not be expected to assimilate systematically to any single Cantonese tone based on its deviant contour shape, if a plausible (though non-ideal) counterpart must be assigned to M4, then C1, would be the most reasonable candidate rather than C4. The tonal assimilations predicted for the Cantonese speakers are thus as illustrated in Figure 2 (below):

TONAL SHAPE	L1 = CANTONESE	L2 = MANDARIN
Level	C1 = High Level	M1 = High Level
	C3 = Mid Level ◄	
	C6 = Low Level	
Rising	C2 = High Rising ◄	M2 = High rising
	C5 = Low Rising	
Dipping		M3 = Low Falling Rising
Falling	C4 = Low Falling ●‥🚫‥	M4 = High Falling

Figure 2. Predicted tonal assimilations by Cantonese learners of Mandarin

Referring to Figure 2, the solid black arrows indicate the mapping of Mandarin tones to the most plausible Cantonese tonal counterpart, projecting

from the cells associated with the relevant Mandarin tones towards the cells of their Cantonese tonal counterparts. Referring to the highlighted cells in Figure 2, the high level Mandarin tone (M1) is expected to map onto, or assimilate to, the mid level Cantonese tone (C3); the high rising Mandarin tone (M2) is expected to map onto the high rising Cantonese tone (C2); and the low falling rising Mandarin tone (M3) is expected to map onto the high rising Cantonese tone (C2). Meanwhile, the dotted bulbous arrow, accompanied by a slash symbol, denotes the idea that the Mandarin high falling tone (M4) is a notably deviant exemplar of the low falling tone (C4) within the Cantonese inventory. Although both M4 and C4 are classified under the 'falling' tonal category, M4 is still not expected to map onto the low falling Cantonese tone in a systematic manner. To emphasize this idea, the typical triangular end of the arrow extending to the C4 cell has been replaced by a circular tip. However, if a Cantonese tonal counterpart had to be chosen for the high falling Mandarin tone outside its typical tonal category, then that counterpart, albeit non-ideal, would most likely be the high level Cantonese tone (C1) because of its comparably high onset frequency.

5. Methodology

5.1 Stimuli Design

Both Cantonese and Mandarin stimuli consisted of isolated monosyllables of the form CV: [ji], [ma], and [wu]. The stimuli were selected such that the onsets of all the target syllables consisted of a nasal or an approximant in order to minimize potential shifts in the tones that could occur with other types of onsets (Khouw & Ciocca 2007) during the production experiments. In addition, the vocalic portion of the syllables intentionally included the cardinal vowels: /i/, /a/, and /u/ because these vowels occupy the most extreme positions of the vocal cavity and are considered areas of spectral salience and stability by Quantal theory (Stevens 1989). The tones manifested on these vowels are thus, by extension, expected to be produced with greater stability than if a different set of vowels were selected. Importantly, /ji/, /ma/, and /wu/, were incorporated because they can combine with the six lexical tones of Cantonese to form actual Cantonese words or morphemes (Kwan *et al*.) and can also combine with the four lexical tones of Mandarin to form actual Mandarin words or morphemes (Oxford University Press 2000). The only exception was the Cantonese syllable /wu/ which can only combine with five of the six primary lexical tones in Cantonese to form actual Cantonese words or morphemes (all tones except the low rising tone, C5, can be combined with the target syllable /wu/ to form real Cantonese lexemes) (Kwan *et al*.). Note that, unlike in Cantonese, the distribution of the four main lexical tones in isolated Mandarin monosyllables can occur with virtually all Mandarin monosyllables with only a few accidental gaps (Howie 1976). Although the use of synthetic stimuli or nonce syllables may have avoided the tonal gap associated with the Cantonese syllable /wu/, authentic and naturally-produced syllables were still used in order to increase

the likelihood that the syllables would sound more natural. Additionally, the segments in these syllables were chosen since they do not favour native Cantonese, native Mandarin, or native English speakers, and the CV form was selected because it does not violate the phonotactic constraints of any of these languages (open syllables are permissible in Cantonese, Mandarin, and English). Note also that, to a certain extent, /ji/, /ma/, and /wu/ formulate real words in English: /ji/ can refer to 'ye'; /ma/ can refer to the informal word for mother: 'ma'; and /wu/ can refer to 'woo'. The only difference is that in English, the manifestation of tones with these syllables does not form contrastive lexical units.

5.2   Cantonese Stimuli

The canonical tonal targets for the Cantonese part of the study were downloaded from the HUMANUM Chinese character database (Kwan *et al*.) since a native Cantonese-speaking Cantonese instructor could not be recruited. The tonal targets obtained from the database were produced by a female instructor of Hong Kong Cantonese whose L1 is Hong Kong Cantonese. Since the duration of each syllabic root in combination with their tones were within approximately half a millisecond of each other, any potential perceptual cue provided by durational differences was considered to be minimal. The intensity for each syllable was set to have a consistent mean intensity of 70 dB. In total, 17 Cantonese monosyllables were used for the Cantonese portions of the study (6 tones with the syllabic root /ji/; 6 tones with the syllabic root /ma/; and 5 tones with the syllabic root /wu/). To minimize the potential interference from adjacent context, these target monosyllables were played in isolation during the various tasks. The audio quality and intelligibility of the Cantonese stimuli were verified by native Cantonese speakers to be good overall with the exception of the target syllable /wu/, which sounded somewhat unnatural.

5.3   Mandarin Stimuli

A male Mandarin instructor who was born and raised in Qiqihar, China and whose L1 is Mandarin was recruited to produce the required target syllables for the Mandarin portion of the study. Target Mandarin syllables were produced in isolation with their canonical tonal shapes such that all the syllables had approximately the same volume and duration. Recordings were made in a sound booth. In total, ten trials, each consisting of the twelve target Mandarin syllables, were recorded. The intensity for each trial was scaled to have a mean intensity of 70 dB using *Praat* 4.5.26 (Boersma & Weenink 1992-2007) in order to reduce the perceptual cue provided by the intensity associated with the Mandarin tones and to be consistent with the intensity level selected for the Cantonese tones. Syllables from each trial were then individually extracted using *Praat* 4.5.26. Optimal tokens were selected such that the tonal shapes had the canonical shapes desired and the duration of the syllables were controlled for. In total, 12 monosyllables were used for the Mandarin portions of the study

(4 tones with the syllabic root /ji/; 4 tones with the syllabic root /ma/; and 4 tones with the syllabic root /wu/). The audio quality and intelligibility of the Mandarin stimuli were verified to be good by a native Mandarin speaker.

5.4 Participants

In total, 23 participants took part in the study: four native Cantonese speakers who previously learned Mandarin as a second language (one female; three male), nine native English speakers (seven female; two male), and ten native Mandarin speakers (8 female; 2 male). Participants were selected based on place of birth, age (of arrival), linguistic background, and education level. All participants ranged from 19 to 28 years of age, all received post-secondary education, and all had sufficient English proficiency to understand the oral instructions provided in English during the study. All participants were healthy at the time of the study, and no hearing or speaking problems were reported by the participants.

5.5 Overview of Tasks

The study involved two perceptual tasks and two production tasks, employing two custom-designed computer programs: *Perceptor* (Kang 2007a) and *Productor* (2007b) for the perceptual and production tasks, respectively. The two perceptual tasks were both 'same-different' AX discrimination tasks, one of which consisted of the target Cantonese tones (160 pairs; 1 trial), and the other of which consisted of the target Mandarin tones (72 pairs; 1 trial). The interstimulus interval for both perceptual tasks was set to 5 seconds and involved a counting task in English. The two production tasks were both shadowing tasks, one of which required the participant to pronounce the pre-recorded Cantonese target syllables after listening to them (17 syllables; 3 trials), and the other of which required the participant to pronounce the pre-recorded Mandarin target syllables after listening to them (12 syllables; 3 trials). Note that the presentation order of tone pairs was taken into account, and tone pairs were counterbalanced. Both the instructors' tone productions served as the standard models to which the productions of the linguistic groups were eventually compared. The Cantonese tone discrimination task was devised in order to determine whether the Cantonese group can perceive the distinctions of their native tones. Meanwhile, the Mandarin tone discrimination task was devised to determine whether there are certain tones that are systematically difficult for Cantonese learners of Mandarin and whether these difficulties are attributable to the influence generated by their native tonal system. The Mandarin tone production task was implemented in order to investigate the production of Mandarin tones by Cantonese learners of Mandarin and to determine the level of perceptual assimilation of these Mandarin tones to their production of Cantonese tones in the Cantonese production task. The production tasks were also included to determine whether the tones which are problematic for Cantonese speakers have a perceptual basis. Comparisons

between the perceptual results and the production results may show correlations. Tones which were difficult to perceive may also be found to be difficult to produce by native Cantonese speakers.

Each of the Cantonese participants performed all four tasks; each of the English participants performed three tasks: the Cantonese and Mandarin tone discrimination tasks and the Mandarin shadowing task; and the Mandarin participants performed two tasks: the Mandarin tone discrimination task and the Mandarin shadowing task. The native English speakers acted as a control group for the discrimination tasks in order to determine the level of perceptual ability of listeners when one's native language is atonal. A basis of comparison could thus be established for the native Cantonese participants' discrimination results to determine whether their native tonal system is advantageous or disadvantageous.

5.6   Perceptual Measurements

Response times (in milliseconds), responses ('Same' or 'Different'), and the accuracy of the responses ('Correct' or 'Incorrect') for each tone pair were automatically recorded in a log text file by *Perceptor* (Kang 2007a) for each perceptual experiment and for each individual participant as applicable.

For each listener, the response times for each tone pair were normalized and ranked according to percentile for subsequent cross-listener analyses. Faster response times were associated with a lower percentile rank and were generally interpreted as being an easier tone pair to discriminate, while slower reaction times were associated with a higher percentile rank and were generally interpreted as being a more difficult tone pair to discriminate.

5.7   Production Measurements

All the target syllables from the Cantonese shadowing task and the Mandarin shadowing task were analysed in *Praat* 4.5.26. The 17 Cantonese target syllables provided by the Cantonese instructor and the 12 Mandarin target syllables provided by the Mandarin instructor were also analysed in *Praat* 4.5.26 in order to isolate the tonal portions from each syllable. Each target tone was selected from the onset of the vowel portion of the syllable with reference to the formants and the amplitude of the waveform.

In total, for the Cantonese shadowing task, 204 tones were extracted for the Cantonese group, while for the Mandarin shadowing task, 360 tones were extracted for the Mandarin control group, 322 tones were extracted for the English control group, and 144 tones were extracted for the Cantonese group. These WAV files were then input into *MATLAB* (The Mathworks, Inc. 2006) along with the WAV files corresponding to the 17 Cantonese tones produced by the Cantonese instructor and the 12 Mandarin tones produced by the Mandarin instructor. Once the data was keyed into *MATLAB*, the program computed the fundamental frequency curves for each input file according to the source codes provided by Kang (2007c). The fundamental frequency contours of each tone

were normalized for each of the speakers according to the formula: $T = [(F_0 - F_{min}) / (F_{max} - F_{min})] \times 5$ (Flynn 2003: 11; Wang *et al.* 2003: 1037), where T is a value ranging from 1 to 5, F_0 is any given point on the fundamental frequency contour, F_{min} is the minimum fundamental frequency value of the speaker's range, and F_{max} is the maximum fundamental frequency value of the speaker's range. For the Cantonese shadowing task, the Cantonese participants' T values for each syllable were compared to the T values of the corresponding syllable produced by the Cantonese instructor, while for the Mandarin shadowing task, the T values for each syllable associated with the tones produced by the Mandarin participants, English participants, and Cantonese participants were compared to the T values of the corresponding syllable produced by the Mandarin instructor. Deviation scores between the participant's productions and the instructors' productions were then calculated. Higher deviation scores correlated with a higher degree of error associated with the participant's tonal production compared to the instructor's production. The deviation scores corresponding to the productions of the four Mandarin tones by each Cantonese participant were then compared in *MATLAB* to the T values obtained for each of the Cantonese tones and the relative errors were derived. By comparing the Mandarin deviation score relative to each of the Cantonese tones, that is, by comparing the production of the Mandarin tones by the Cantonese participant to their productions of their native Cantonese tones, the closest Cantonese tonal counterpart for each Mandarin tone could be quantitatively determined. The lower the deviation between the Mandarin tones and the Cantonese tones (as produced by the Cantonese participant), the more likely these tones matched each other. The perceptual assimilation patterns could thus be analysed by comparing the relative differences between the productions of the non-native tones versus the productions of the native tones. For instance, if the high level Mandarin tone was compared to each of the Cantonese tones, and the lowest error found was between the high level Mandarin tone and the high level Cantonese tone, then, that would suggest that the high level Mandarin tone would most likely map onto the high level Cantonese tone. Another way of analysing this situation would be that the high level Mandarin tone is produced by a speaker such that it is most acoustically similar to their production of the high level Cantonese tone.

6.   Results & Discussion

6.1   Cantonese Tone Perception

Referring to Table 3, results for the Cantonese tone perception task showed that C2 and C5 (the two rising tones) were the most difficult for the Cantonese speakers to discriminate, while the English speakers had difficulty with a wide range of tones. The Cantonese speakers' results are consistent with previous studies (e.g., Fok 1974) in which the two rising tones were found to be easily confused. Overall, the Cantonese speakers outperformed the English speakers, showing that the Cantonese speakers' native tonal inventory played a facilitative

role in their perception of Cantonese tone and verifying that they can indeed perceive their own native tones. If the results had shown that Cantonese speakers were unable to discriminate their native tones accurately, then the posited assimilation patterns would have no basis and subsequent analyses would have had to be abandoned.

Table 3. Results for Cantonese tone perception

	CANTONESE	ENGLISH
Difficult tones	C2, C5	Wide range
Avg. # Errors	4	13
Range of Errors	0–6	4–32

6.2 Mandarin Tone Perception

Referring to Table 4, results for the Mandarin tone perception task showed that native Mandarin speakers had the most difficulty with tone pairs involving M2 (high rising tone), while the Cantonese speakers had the most difficulty with tone pairs involving M4 (high falling tone).

Table 4. Results for Mandarin tone perception

	MANDARIN	CANTONESE
Difficult tones	M2	M4 (M4–M4; M2–M2; M3–M1)
Avg. # Errors	2	2
Range of Errors	0–4	0–3

Cantonese participants' difficulty with M4 can be attributed to the absence of a Cantonese tonal counterpart for the high falling tone. In particular, the pairs M4-M4, M2-M2, and M3-M1 were difficult for the Cantonese speakers to discriminate. The M4-M4 pair may have been difficult because the Cantonese speaker potentially perceived one of the falling tones as a high level tone (perceptual interference due to native tonal inventory). Similarly, the M2-M2 pair may have been misperceived as two different tones since there are two different rising tones in the Cantonese tonal inventory. Finally, the M3-M1 pair may have been difficult for the Cantonese speakers because there is no dipping tone in Cantonese, so the final portion of the dipping tone may have been stored in memory as merely a high level tone (rather than the posited C2 tone), thereby resulting in the perception of two pairs of high level tones rather than a pair comprised of two different tonal contours (thus, the pair was perceived as the same when they are actually different). Nevertheless, the errors generated by the Mandarin and Cantonese groups were still comparable overall (both averaging 2 errors). In fact, few errors were generated across all linguistic groups presumably because of the psychoacoustic salience of Mandarin tones. Overall, the Mandarin tone perception by the Cantonese speakers was as

accurate as the Mandarin control group. For the Cantonese group, the difficulty of tones tended to follow the pattern: M2, M3 > M4 > M1, where the high rising and low falling rising tones were the most difficult, followed by the high falling and high level tones, respectively.

6.3   Cantonese Tone Production

The Cantonese tone production task showed that the Cantonese speakers had the most difficulty with the two rising tones, C2 and C5. This result is not surprising considering that these two tones have similar contour shapes and considering that the Cantonese speakers had difficulty discriminating these two tones during the Cantonese perception task. In this case, tonal production seems to have a perceptual basis: C2 and C5 were difficult to perceive and thus difficult to produce by the Cantonese speakers. Overall, the Cantonese productions by the Cantonese speakers were deemed acceptable for comparison with their Mandarin tone productions since comparisons with the Cantonese instructor's tonal productions showed low average deviation scores (0.5 value).

6.4   Mandarin Tone Production

Results for the Mandarin tone production task showed that the Mandarin speakers had the lowest deviation scores, followed by the Cantonese speakers and English speakers, respectively. This result shows that native speakers still outperform non-native speakers' tone productions, although the existence of a native tonal inventory can be helpful in non-native tonal production.

Finally, the tonal assimilation patterns demonstrated by the Cantonese participants are displayed in Figure 3. Overall, the assimilation patterns posited in the hypothesis section were confirmed: M2 mapped to C2, M3 mapped to C2, and M4 mapped to C1 (Note that a dotted line extends from M4 to C1 in Figure 3 because C1 (a level tone) is outside the tonal category of M4 (a falling tone); however, M1 assimilated to C1 most frequently instead of to C3 as previously posited. There are two possible reasons for this mapping: either the predicted perceptual mapping of M1 onto C3 is the most plausible one and the small sample size was not conducive for demonstrating this assimilation pattern, or the high level Mandarin tone may have been produced higher because there is only one level tone in Mandarin. Upon hearing a high level tone, then, the Cantonese speakers simply produced a high tone coinciding with the acoustic tonal space occupied by their high level tone instead of a lower one, such as the mid-tone, which is the most acoustically similar to the Mandarin high tone. Shifts in tone height corresponding to the high level Mandarin tone would thus result in a level tone that would still not be confounded with any of the other Mandarin tones because the other tones have such distinct shapes. It would seem, then, that M1 is the best exemplar of C1 rather than C3 because the focus is on producing the M1 tone as high.

Figure 3. Resultant tonal assimilation patterns

7.   Conclusion

Based on the results obtained from both the Mandarin tone perception and production tasks, more insight has been gained regarding the potential patterns of tonal assimilation that Cantonese speakers undergo during Mandarin tone perception and production.

Although the number of participants in this study did not lend itself to make conclusive claims about the extension of PAM (Best 1995) to suprasemegmentals, the study still strived to fill the gap in the literature regarding the non-native perception of tones by L2 learners. Importantly, the results suggest that even native Cantonese speakers who are L2 learners of Mandarin still have perceptual difficulty differentiating familiar non-native tones. That is, the contribution of the L1 to the perception of L2 tones still persists regardless of training in the target language. The study also reaffirmed that listeners with L1 experience are superior to naïve listeners for certain tasks (e.g., in the Cantonese perception task, the Cantonese group outperformed the English group). Results may vary, however, depending on the type of task implemented. Finally, an important contribution of the current study is that it showed that the basic perceptual assimilation hypotheses outlined in PAM (Best 1995), which is a model for non-native segmental perception, can be used to model the perception of lexical tones. However, PAM (Best 1995) is overall quite general in its predictions and is predominantly a perceptual model. In the future, more quantitative models need to be devised to reinforce the predictions postulated by PAM (Best 1995) and to link production studies to this perception model. The current study, which used both perception and production tasks, may thus act as a guide to developing more robust experimental procedures for modelling tone assimilation patterns from a non-native system to a native system. Future work could also incorporate Mandarin L2 learners of Cantonese to investigate the perceptual assimilation of native tones from a smaller tonal inventory to a larger tonal inventory, thereby complementing this study.

References

Best, Catherine T. 1995. A direct realist perspective on cross-language speech perception. In Winifred Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*. (pp. 167-200). Timonium, MD: York Press.

Blicher, Deborah L., Randy L. Diehl, and Leslie B. Cohen. 1990. Effects of syllable duration on the perception of the Mandarin Tone 2/Tone 3 distinction: evidence of auditory enhancement. *Journal of Phonetics*, *18*, 37-49.

Boersma, Paul, and David Weenink. 1992-2007. *Praat* 4.5.26. Computer program.

Chao, Yuen-Ren. 1947. *Cantonese primer*. Cambridge, MA: Harvard-Yenching Institute, Harvard University Press.

Fok Chan, Yuen-Yuen. 1974. *A perceptual study of tones in Cantonese*. Hong Kong: Centre of Asian Studies Publications - University of Hong Kong.

Flynn, Choi-Yeung-Chang. 2003. *Intonation in Cantonese*. München: LINCOM Europa.

Hallé, Pierre A., Yueh-Chin Chang, and Catherine T. Best. 2003. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*(3), 395-421.

Howie, John M. 1976. *Acoustical studies of Mandarin vowels and tones*. Cambridge, UK: Cambridge University Press.

Kang, Jeffrey J. 2007a. *Perceptor*. Computer program.

Kang, Jeffrey J. 2007b. *Productor*. Computer program.

Kang, Jeffrey J. 2007c. Source codes for *MATLAB* analyses.

Kwan, Tze-Wan, Wai Sang Tang, Tze Ming Chiu, Leonardo L. Y. Wong, Denise Wong, and Li Zhong. HUMANUM Chinese character database: *A Chinese Character Database: With word-formations phonologically disambiguated according to the Cantonese dialect*. Cantonese tones retrieved December, 2006 from: http://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/.

Khouw, Edward, and Valter Ciocca. 2007. Perceptual correlates of Cantonese tones. *Journal of Phonetics*, *35*(1), 104-117.

Lee, Yuh-Shiow, Douglas A. Vakoch, and Lee H. Wurm. 1996. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, *25*(5), 527-542.

Oxford University Press. 2000. *Concise English-Chinese Chinese-English dictionary*. Hong Kong: Oxford University Press (China) Ltd.

Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics*, *17*, 3-45.

The Mathworks, Inc. 2006. *MATLAB*. Computer program.

Wang, Yue, Allard Jongman, and Joan A. Sereno. 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, *113*, 1033-1043.

Whalen, Douglas H., and Yi Xu. 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, *49*, 25-47.