Perception of Mandarin Tones by Mandarin and English Listeners*

Xianghua Wu[1]
xianghua_wu@sfu.ca

Hua Lin[2]
hualin@uvic.ca

[1]Department of Linguistics, Simon Fraser University, Canada
[2]Department of Linguistics, University of Victoria, Canada

Abstract

This paper investigated the perception of Mandarin tones by Mandarin and English speakers; in particular, whether the perception is categorical or not. Twelve native Mandarin speakers from China and twelve native Canadian English speakers with no prior experience in tone language took a discrimination test and an identification test. Created using the Mandarin syllable *shi* in the four Mandarin tones, the test stimuli consisted of two tonal continua, one between Tone 1 and Tone 4 and the other between Tone 2 and Tone 3. Each continuum consisted of eight tokens. The results indicated that the English listeners' perception of Mandarin tones was not categorical whereas the Mandarin listeners' was, but only between the Tone 1/Tone 4 contrast, and not between the Tone 2/Tone 3 contrast. Nevertheless, the English listeners showed strong sensitivity to the Mandarin tone contrasts. The results were discussed in the framework of Best's Perceptual Assimilation Model (Best 1994 and 1995).

Keywords

tone, tone perception, categorical perception, acoustic phonetics, Mandarin Chinese

---

* We wish to acknowledge the partial support for this study from the project 'the Formal Syntactic Theory and A Study of Some Essential Chinese Syntactic Issues' (05JJD740184) from the Key Project Schema of the Advanced Research Bases in Humanities and Social Sciences, the Ministry of Education, China. Thanks also go to the anonymous reviewers of this paper.

1 Introduction

Tone and non-tone languages differ in the use of pitch patterns at the lexical level. Some tone languages, such as Mandarin, employ pitch heights and direction (contours) on the syllable to differentiate word meanings. For instance, syllable *ma* in a high level pitch means *mother*, while with a high-to-low falling contour means *to blame*. Non-tone languages, such as English, on the other hand, use pitch to indicate, among other matters, different syntactic categories, phrasal intonation and meaning contrast in words. For instance, the two disyllabic words in the identical segmental form *forbear* can be realized as FOREbear and forBEAR, the former being a verb meaning *to refrain from* with the accent on the first syllable and the latter a noun meaning *ancestor* with the accent on the second syllable. As is often the case, the pitch patterns usually extend over at least two syllables in English.

Such use of word-level prosody has been researched on in many previous studies, an example of which is the dichotic listening (CL) research. DL studies are interested in whether there is a hemispheric bias in the perception of word-level prosody by native listeners. DL studies on tone perception provided evidence supporting a left hemispheric advantage for native listeners of tone languages processing lexical tones (Moen 1993, Van Lancker and Fromkin 1973, and Wang *et al.* 2004) and for native English listeners processing English word-level stress (Arciuli, J., & Slowiaczek 2007). These studies also revealed a distinctive lateralization pattern for non-native listeners compared to the native listeners. However, in Van Lancker and Fromkin's (1973) study, English listeners showed no hemispheric preference when processing Thai tones. The same result was obtained in Wang *et al.* (2004) for the Norwegian and English listeners processing Mandarin tones. These findings suggest that on the one hand, the perceptual patterns for word-level prosody are affected by native language background; on the other hand, the perception of non-native prosody is not completely unattainable by the non-native listeners.

Another approach to word-level prosody examines the *categorical perception* of this type of prosody. The classic description of a categorical perception of speech refers to the phenomenon where small steps along an acoustic continuum will produce perceivable differences when they occur between phonetic categories, but not when they occur within a phonetic category (Gandour 1978). It is suggested that the native listeners are able to distinguish a phonemic contrast categorically whereas non-native listeners are not. For example, in the perception of word-level prosody such as lexical tones, some research indicates that only native listeners show categorical perception whereas non-native listeners do not (e.g., Hallé *et al.* 2004, Wang 1976, and Xu *et al.* 2006). Nevertheless, these studies also reveal that linguistic experience is not the only factor that affects categorical perception. The nature of the acoustic cues, memory retention and other psychophysical factors also play significant roles in the categorical perception of lexical tones. With regards to native listeners, Abramson (1977) and Francis *et al.* (2003) find that Thai and Cantonese listeners show non-categorical perception of the level tone contrasts in their native languages. Shen and Lin (1990) claim that partial loss of perceptual cues is likely to cause native speakers to lose categorical perception of the contrast between the Mandarin rising tone (Tone 2) and falling-rising tone (Tone 3). For the non-native listeners, even though a categorical perception is not obtained, these studies argue that substantial sensitivity to lexical tones do exist, though to a lesser extent than in the native listening capacity.

This study was carried out based on two gaps existing in the previous categorical perception studies: (1) Few studies compared the performance of Mandarin and English listeners of the Mandarin tone contrasts. (2) Few studies examined closely the perception of the contrasts between Tone 2/Tone 3 and Tone 1/Tone 4, even though these contrasts are known to be the most confusing in Mandarin (Blicher *et al.* 1990, Huang 2001, Kiriloff 1969, Li and Thompson 1977, and Sun 1998). The present study compares Mandarin and English listeners' perception of the two Mandarin tonal contrasts; It aims to find out what factors substantially affect categorical perception of these Mandarin tone contrasts by both native and non-native speakers.

2 Methodology

2.1 Participants

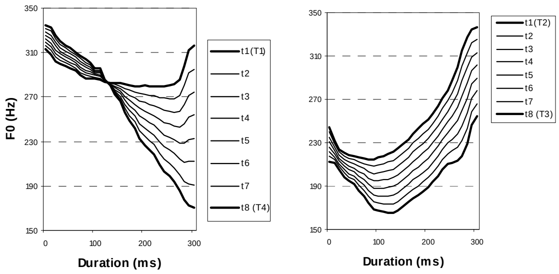
A total of 24 subjects, 12 native Mandarin and 12 native English speakers, participated in this study. The Mandarin subjects, aged 22-38, included three males and nine females, all students at the University of Victoria. None had been exposed to other tone languages or had received formal musical training. All spoke English as a second language. The English subjects, aged 22-46, consisted of eight males and four females. None had any previous experience with tone languages. Eleven were students at the University of Victoria and one worked in Victoria. Six of the 12 reported having received formal musical training. All the 24 participants from the two groups reported normal speech and hearing.

2.2 Stimuli

Stimuli were created from four naturally produced Mandarin tones on the target syllable *shi*. They represent four commonly used Mandarin words such 师 "teacher" in Tone 1, 十

"ten" in Tone 2,使 "to let" in Tone 3, and 是 "to be" in Tone 4, respectively. The four words were read by a female native Mandarin speaker and were recorded.

Two tone continua, one ranging from Tone 1 to Tone 4 (high level to high falling), and the other from Tone 2 to Tone 3 (rising to falling-rising) were created using the recorded materials. Each tonal continuum consisted of eight tokens with six interpolated by equal steps between the two end points. All the tokens on each continuum had the same amplitude envelope and duration. The speech analysis software Praat (Version 4.2.31) was used to equalize the duration and amplitude of the syllables (see Fig. 1[1]).



A B

Figure 1. The eight tokens (labelled as *t1*, *t2*, *t3* …, etc.) for the syllable *shi* in the two continua: (A) on the left is the T1-T4 continuum; (B) on the right is the T2-T3 continuum.

2.3 Procedure

An AX (same/different) discrimination test and an AXB (forced-choice) identification test were created using the above mentioned tokens. All 24 subjects participated in the these tests. An AXB identification task was chosen over the labelling task more commonly adopted in categorical perception experiments because a labelling task would be inappropriate for non-native subjects who were not acquainted with Mandarin tones (Hallé *et al.* 2004).

The experiment was conducted in the Phonetics Laboratory at the University of Victoria. Each subject was tested individually on a PC computer. They first did an AXB forced-choice identification task in which they answered whether the second token (token "X") in a triplet sounded more like the first token (token "A") or the third token (token "B"), and then an AX discrimination task in which they chose between "same" or "different" for

[1] In what follows, we will use T1, T2, T3, and T4 interchangeably with Tones 1, 2, 3 and 4, and t1, t2, … to represent the steps in the continua.

each presented stimulus pair. Each task contained seven randomized blocks, twenty triplets or pairs in each block. The instructions and answer sheets were presented in Chinese to the Chinese listeners and in English to the English listeners. To avoid fatigue, the listeners took a short break between the identification and discrimination tasks. The whole session lasted 40 minutes.

2.4 Data analysis

Two sets of measures were computed for each listener: (1) the intercept and the slope of the category boundaries in the identification test; and (2) the discrimination accuracy and location of the peak points in the discrimination test.

The results of the identification tests were examined through narrow-range PROBIT analysis (Best and Strange 1992) which fits the (negative) cumulative normal distribution to the curve defined by the probabilities (percent response) versus the tokens. A weighted regression was used so that probabilities close to 0 and 1 have less effect on the fitted line (Mackain *et al.* 1981, Hallé *et al.* 2004, and Lesperance, personal communication). According to this analysis, the category boundary refers to the 50% intercept of the best-fitted three-point ogives, the three points closest to the 50% crossover on the identification curves. The slopes of these ogives (1/SD) indicate “the peak rate of change in category labelling at the crossover, and were used as a reflection of the steepness of the category boundaries, i.e., larger values indicate steeper functions” (6:313). To choose the three best points which were closest to the 50% crossover, a general principle of “two above and one below” was used in this study (Hallé, personal communication).

The discrimination results were examined from three perspectives: (1) the average percentage of correct response across all the tokens; (2) the existence of one or two adjacent peak points which reach the highest discrimination level; and (3) the correspondence of the location of peak point(s) and the identification intercept.

3 Results

3.1 Identification task

As seen in Figure 2, the identification curves exhibit sigmoid shapes and steep slopes on both continua across the two groups, which was likely the result of using an AXB identification task (Hallé *et al.* 2004).

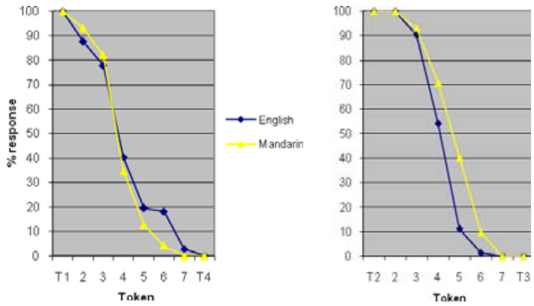
Although the data for both groups yielded sigmoid-shaped curves, compared to the English group, the Mandarin group showed a smoother descending curve from left to right for the T1-T4 continuum (Figure 2A), whereas the English group identified token 6 as token 1 more often than did the Mandarin group. The figure was 18.06% for the English group, but only 4.17% for the Mandarin group. In other words, token 6 sounded more like T4 to the Mandarin group than to the English group.

The identification data presented in Table 1 show category boundaries falling on the left of the centre point (4.5 for an eight-token continuum) for the T1-T4 continuum, meaning that more tokens across the boundary were identified as the endpoint T4. The data also indicate that the intercept falls farther away from the centre point for the Mandarin group (3.72) and closer to the centre point (3.79) for the English group. The difference between the groups, however, is not significant ($F (1, 10) = 0.028$, $p = 0.87$). The slope data for the T1-T4 continuum reveals to some extent the difference between the native and non-native groups. Although an ANOVA did not show a significant difference between the two language groups at the 0.05 significance level ($F (1, 10) = 1.408$, $p=0.263$), it is readily apparent that the Mandarin group has a steeper slope than the English groups (1.64 vs. 1.24), suggesting that the identification function is higher for the native Mandarin speakers.

The results for the identification of the T2-T3 continuum are shown in Figure 2B and Table 1. Figure 2B does not exhibit any difference in the shape of the identification curve among the groups. The locations of the intercept, however, are not as convergent as those shown in Figure 2A. The data are given in Table 2: Mandarin, 4.74, and English, 4.11. That is, the intercept of the category boundary falls to the right of the center point (4.5 for an eight-token continuum) for the Mandarin group, while it falls to the left of the centre point for the English groups.

The locations of the intercept indicate that the English listeners were more likely to identify the tokens as T3, whereas the intercepts tended to cluster around the centre point for the Mandarin group. ANOVA shows this result was marginally statistically significant at the 0.05 level ($F (1, 10) = 4.279$, $p = 0.065$).

The slopes of the boundaries by token number are also given in Table 1: Mandarin 1.32 and English 1.51. The English group exhibited a steeper slope than the Mandarin group. These data suggest that the native Mandarin listeners did not have much of an advantage over the non-native listeners in identifying these two tones. Further ANOVAs performed on the steepness of the slopes failed to show any significance on the effect of the continuum between the two groups. However, relative to the slope for the T1-T4 continuum, there was a clear tendency for the identification function to decline for the Mandarin group (from 1.64 to 1.32) along the T2-T3 continuum, and to increase for the English group (from 1.24 to 1.51). This pattern suggests that for the Mandarin group, T1 and T4 are more readily distinguishable than T2 and T3, but the reverse is the case for the English group.



A B

Figure 2. Identification curves for the Mandarin, and English groups: (A) on the T1-T4 continuum; and (B) on the T2-T3 continuum.

Continuum	Intercept (token number)		Slope (1/SD value)	
	Mandarin	English	Mandarin	English
T1-T4	3.72	3.79	1.64	1.24
T2-T3	4.74	4.11	1.32	1.51

Table 1. Intercept and slope results

3.2 Discrimination task

The data on the discrimination tests (see Table 2) do not show any advantage for the Mandarin group over the English group. For both continua, the English group out-performed the Mandarin group. Further examination was carried to examine performance of the English subjects with musical training in order to determine whether the high discrimination level exhibited by the English group resulted from the participation of musically trained subjects.

Continuum	**Mean (SD)**	
	Mandarin	English
T1-T4	70.92 (0.14)	73.28 (0.16)
T2-T3	61.89 (0.1)	71.5 (0.13)

Table 2. Mean correct discrimination scores (%) from the AX discrimination task

As shown in Table 3, the musicians outperformed the non-musicians significantly in the discrimination test ($F (1, 34) = 14.85$, $p<0.001$ and $F (1, 34) = 12.14$, $p=0.001$ for T1-T4 and T2-T3 continuum, respectively). A closer look at the data reveals that the higher discrimination levels of the English group is attributable to the performance of the musicians. According to the data presented in Tables 2 and 3, Mandarin listeners outperformed English language speakers who have no musical training (70.92% vs. 67%, 61.89% vs. 57% for the T1-T4 and T2-T3 continua, respectively). These data suggest that L1 tone experience may be the crucial factor to affect tone discrimination for listeners who have had no musical training.

English	T1-T4 Mean (SD)	T2-T3 Mean (SD)
Pooled	73.28(0.16)	71.5(0.13)
Music	85.83(0.17)	74.4(0.12)
Non-music	67(0.12)	57(0.08)

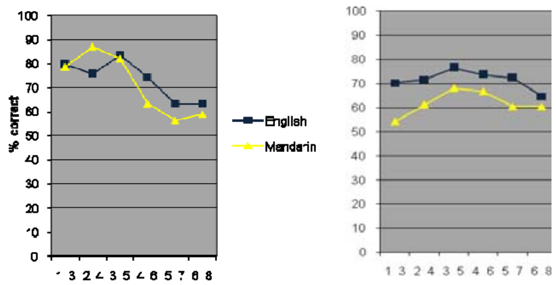
Table 3. Mean correct scores (%) for the English group and its subgroups on the AX discrimination task.

In addition to group effect, tone continuum effect was also statistically examined with ANOVA. As shown in Table 2, the Mandarin group performed better in discriminating the T1-T4 than the T2-T3 continuum, and this difference was significant ($F (1, 70) = 9.945$, $p = 0.002$). However, despite its high performance scores, the continuum effect was not significant ($F (1, 70) = 0.264$, $p = 0.609$) for the English group.

Figure 3 displays the discrimination peaks on the curves. A discrimination curve for a two-category continuum usually has one peak point or two continuous peak points (Hallé *et al.* 2004). An ANOVA show that neither of the peak points on the discrimination curves for the English group represent a significantly higher level of discrimination than any of the other stimulus pairs ($F (5, 30) = 1.76$, $p= 0.153$; and $F (5, 30) =0.52$, $p = 0.76$ for the T1-T4 and T2-T3 continua, respectively).

The Mandarin group showed a discrepancy in the discrimination of the native tone contrasts T1-T4 and T2-T3. Although both discrimination curves show peak points, these discrimination peaks are only significant for the T1-T4 continuum ($F (5, 30) = 14.974$, $p < 0.0001$), not for the T2-T3 continuum ($F (5, 30) = 1.923$, $p = 0.12$). These results obtained for T2-T3 are counter-intuitive, given that lexical tone phonemically differentiates word meanings for native Mandarin speakers.

The last question concerns the correspondence between the locations of the peak points and the identification intercepts. As shown in Table 1, the locations of the intercepts on the identification curves for the T1-T4 continuum are around 3.72 and 3.76 (in stimulus number) for the Mandarin group. The peak points fall between stimulus pairs 2-4 (87%) and 3-5 (82%) for the Mandarin group. In other word, the identification intercepts and discrimination peaks tend to coincide.



A B

Figure 3. Two-step discrimination curves for the Mandarin, and English groups: (A) on the T1-T4 continuum; and (B) on the T2-T3 continuum

4 Results

The results of this study show different patterns in the perception of Mandarin tones by native and non-native listeners. For the native listeners, as expected, perception of Tone 1/Tone 4 contrast is categorical. A steep slope was obtained on the identification curve and the discrimination peak reached significantly higher level compared to other tokens. And, the position of the peak coincided with the intercept on the identification curve. The perception of Tone 2/Tone 3 contrast, however, did not yield comparable result for the native listeners. More on this shortly. For the English listeners, the criteria for a categorical perception were not satisfied; thus their perception of Mandarin tones is non-categorical.

4.1 Results for Mandarin speakers

The results of this study revealed a difference in the Mandarin group’s perception of the Tone 1/Tone 4 and Tone 2/Tone 3 contrasts. The difference suggests that despite the fact that all the four tones were selected in the same way, the similarities between Tones 2 and 3 may have created confusion not only for the non-native listeners but also for the native listeners. The results are not too surprising given that Tones 2 and 3 have been identified in a number of studies as one of the most confusing tone pairs in Mandarin (Blicher et al. 1990, Huang 2001, Kiriloff 1969, Li and Thompson 1977, and Sun 1998). Apparently, the confusion is due to the similarities in the occurrence of the turning point and the F0 difference between the tonal onset and the turning point. Shen and Lin (1990) suggest that the timing of the turning point and the F0 difference between the tonal onset and the turning point have to meet a threshold; that is, the turning points are at 14% and 48% of the total duration for Tone 2 and Tone 3, respectively, while the F0 difference between the tonal onset and the turning point is 17.5 Hz for Tone 2 and 38.6 Hz for Tone 3 (p. 152). Therefore, those tones with the turning point occurring between 14% and 48%, and the F0 difference which ranges from 17.5 to 38.6Hz between the tonal onset and the turning point is most likely to cause confusion. In this study, both Tones 2 and 3 fall into the confusion range: the turning point was at 27% for Tone 2 and 40% for Tone 3 of the total pitch length, whereas the F0 between the tonal onset and the turning point was 29.8 Hz for Tone 2 and 47.25 Hz for Tone 3. Thus it is clear that because of the loss of certain perceptual cues due partially to the fact that the tokens for the two tones were synthesized using the same duration, the tokens on the Tone2/Tone3 continuum in this study were difficult for the native Mandarin speakers to identify and discriminate accurately.

4.2 Results for English speakers

As expected, English listeners perceived both Mandarin tonal contrasts non-categorically. Nevertheless, as observed by Hallé *et al*. (2004), speakers of non-tone languages have substantial sensitivity when perceiving lexical tones. This is confirmed in this study in which the identification curves exhibit sigmoid shapes and steep slopes on both continua not only for the Mandarin group but also for the English group. Moreover, the English listeners with prior music training exhibited significantly higher advantage in the discrimination of Mandarin tone contrasts.

4.3 Perceptual Assimilation Model (PAM)

The results in this study have been examined in the context of the Perceptual Assimilation Model (PAM) developed by Best and her colleagues (Best 1994 and 1995, and Best *et al.* 1988). These studies attempted to account for perception patterns in terms of assimilation caused by the correspondence between native and non-native phonetic categories. Two types of correspondences in the revised version of PAM are of concern here: (1) Both Uncategorizable (UU) and (2) Uncategorized versus Categorized (UC). In the UU pattern, both non-native sounds fall within the phonetic space outside of any particular native category and can vary in their discriminability as uncategorizable speech sounds. Hallé *et al.*

(2004) and So (2005) classified tone assimilation as belonging to the UU pattern for native speakers of non-tone languages (French and English, respectively). In their studies of Mandarin tone perception, speakers of French or English performed worse than native Mandarin speakers and native speakers of other tone languages. Nevertheless, the French and English speakers were not completely insensitive to tone contrasts. For instance, while the French listeners failed to categorize any of the Mandarin tones presented in the identification and discrimination tests, they did exhibit psychophysical sensitivity to the Mandarin tones (Hallé *et al.* 2004). Similar result is also seen in the performance of some English listeners in the present study. The English subjects some of whom are musicians outperformed even the native Mandarin listeners in tone discrimination, despite the absence of tonal distinctions in their L1. Their performance on tone perception ranged from fair to good depending on the perceived salience of the phonetic differences involved, as predicted by PAM for the UU assimilation pattern.

5 Discussion

This paper examined the perception of two Mandarin tone contrasts by speakers who differ in their experience with linguistic tone. Native language background was found to be one of the most influential factors in tone categorization. The Mandarin listeners outperformed the English listeners in the categorization of the Mandarin tones, especially with the T1/T4 contrast. The English listeners, on the other hand, failed to categorize the Mandarin tone contrasts despite the presence of high discrimination levels.

Mandarin and English listeners both perceived the T2/T3 contrast non-categorically. The inability of the Mandarin speakers to categorize this native tone contrast may have been caused by the partial loss of perceptual cues. According to Shen and Lin (1990) and Blicher *et al.* (1990), the period from the tonal onset to the turning point and the location of the turning point are crucial cues that distinguish T2 and T3. While the synthesized T2 and T3 used in this study retain the majority of the F0 information available in natural speech, the critical cues noted above were not present.

The English speakers' lack of L1 tone experience is the most likely explanation for their failure to make a categorical distinction between the two tonal contrasts. Musical experience was found to affect the performance of the English participants in the discrimination task; participants with musical training significantly outperformed those without musical training. However, there was no significant difference between the location of the category boundary and the steepness of the boundary, which suggests that despite their good performance in discrimination, the English listeners with a musical background still perceived lexical tones non-categorically. Nevertheless, the English listeners were not insensitive to the Mandarin tone contrasts, although their perception of the lexical tones was most likely psychophysically based. The results of this study support the Perceptual Assimilation Model (Best 1994 and 1995, and Best *et al.* 1988). For the English listeners, lexical tone contrasts may be perceived in the manner of “both uncategorized” (UU) pattern of assimilation.

Needless to say, this study has limitations. One most obvious is that none of the Mandarin subjects had musical training background. Future studies should endeavor to recruit subjects with and without musical training so that comparison can be made to see if native musicians would outperform non-native musicians on the discrimination and identification tasks.

References

Abramson, A. S. 1977. The noncategorical perception of tone categories in Thai. *Haskins Laboratories Status Report on Speech Research*, vol. 51-52, pp. 91-100.

Arciuli, J., & Slowiaczek, L.M. 2007. The where and when of linguistic word-level prosody. Neuropsychologia, vol. 45, pp. 2638-2642.

Best, C. T. 1994. The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model. In Goodman, J. C. and H. C. Nusbaum (eds.) *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, pp. 167-224. MIT Press, Cambridge, MA.

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, pp. 171-204. Timonium, MD: York Press, Inc.

Best, C. T., McRoberts, G. W., & Sithole, N. M. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, vol. 4, pp. 45-60.

Best, C. & Strange, W. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, vol. 20, pp. 305-330.

Blicher, D. L., Diehl, R. L., & Cohen, L. B. 1990. Effects of syllable duration on the perception of the mandarin tone 2/Tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, vol. 18, no. 1, pp. 37-49.

Francis, A. L., Ciocca, V., & Ng, B. K. C. 2003. On the (non)categorical perception of lexical tones. *Perception & Psychophysics*, vol. 65, no. 7, pp. 1029-1044.

Gandour, J. 1978. The perception of tone. In Fromkin, V. A. (ed.), *Tone: A Linguistic Survey*, pp. 41-76. New York: Academic Press.

Hallé, P. A., Chang, Y., & Best, C. T. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, vol. 32, no. 3, pp. 395-421.

Huang, T. 2001. The interplay of perception and phonology in Tone 3 sandhi in Chinese Putonghua. In E. Hume and K. Johnson (eds.) *Studies on the Interplay of Speech Perception and Phonology, Ohio State University working papers in Linguistics*, vol. 55, pp. 23-42.

Kiriloff, C. 1969. On the auditory perception of tones in Mandarin. *Phonetica*, 20, 63-67.

Li, C. N. & Thompson S. A. 1977. The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, vol. 4, pp. 185-200.

Mackain, K.S, Best, C.T. & Strange, W. 1981. Distinguishing universal and language-dependant levels of speech perception: evidence from Japanese listener's perception of English, *Applied Psycholinguistics*, vol. 2, pp. 369-90.

Moen, I. 1993. Functional lateralization of the perception of Norwegian word tones--evidence from a dichotic listening experiment. *Brain and Language*, vol. 44, no. 4, pp. 400-413.

Shen, X. S. and Lin, M. 1990. A Perceptual Study of Mandarin Tones 2 and 3. *Language and Speech*, vol. 34, nov. 2, pp. 145-156.

So, C. 2005. *How do L1 prosodic backgrounds influence Cantonese and English speakers to perceive Mandarin tones?* 1st ASA Workshop on Second Language Speech Learning, Vancouver, Canada. Manuscript.

Sun, S. H. 1998. *The development of a lexical tone phonology in American adult learners of standard mandarin Chinese*. Ph.D Dissertation. Hawaii University.

Van Lancker, D. and Fromkin, V.A. 1973. Hemispheric specialization for pitch and "tone": Evidence from Thai. *Journal of Phonetics*, vol.1, pp. 101-109.

Wang, W.S.-Y. 1976. Language Change. In S.R. Harnad, H.D.Stek-lis, & J. Lancaster (eds.), *Origins and evaluation of language and speech, Annals of the New York Academy of Sciences*, vol. 280, pp. 61-72.

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. 2004. The Role of Linguistic Experience in the Hemispheric Processing of Lexical Tone. *Applied Psycholinguistics*, vol. 25, no. 3, pp. 449-466.

Xu, Y., Gandour, J., & Francis, A. 2006. Effects of language experience and stimulus complexity on categorical perception of pitch direction. *Journal of the Acoustical Society of America*, vol. 120, no. 2, pp. 1063-1074.