Perception of non-native tonal contrasts: Effects of native phonological and phonetic influences

Connie K. So

MARCS Auditory Laboratories
University of Western Sydney, Australia
c.so@uws.edu.au

Abstract

This study examined the perception of Mandarin tones by two groups of Cantonese and Japanese (naïve) listeners. An identification task was given and their responses were analyzed in terms of A-prime scores and tonal errors. The results indicated that the performance of the Cantonese listeners was compatible with that of the Japanese listeners in A-prime scores and tonal errors. The listeners' tonal errors also showed that both the listener groups made considerable amount of errors for the T1-T4 and T2-T3 pairs, but the Cantonese listeners made more noticeable errors for the T1-T4 and T2-T3 pairs than did the Japanese listeners. The discrepancies in the performance between the two listener groups could be explained in the framework of the Perceptual Assimilation Model (PAM: Best, 1994, 1995). The findings imply that linguistic experience of tones does not necessarily facilitate the perception of non-native tones. However, the phonemic status and the phonetic similarities or dissimilarities between the prosodic systems of the target and native languages play a more important role in the perception of non-native tone contrasts. In addition, the results revealed some previously unnoticeable asymmetrical patterns for Mandarin tone perception by non-native listeners. Both the listener groups exhibited great confusion about the tones in the T1-T2, T1-T4, and T2-T3 pairs, but they made apparently fewer errors for the other three pairs of tones, T1-T3, T2-T4, and T3-T4. These imply that non-native tonal perception is also influenced by the phonetic characteristics of target tones.

1. Introduction

Adults have difficulties discerning non-native segmental distinctions that are not contrastive in their own languages. For example, Japanese listeners have difficulties perceiving the English approximants /r/ and /l/ (e.g., MacKain, Best, & Strange, 1981; Logan, Lively, & Pisoni, 1991). Canadian French listeners have problems in perceiving the English fricatives, /ð/ and /θ/ (Jamieson & Morosan, 1986, 1989). Similarly, native English speakers find it hard to discern the differences between certain non-native sound segments, such as the Hindi voiceless dental-retroflex stops (Werker, Gilbert, Humphrey, & Tees, 1981), and the French front rounded vowels /y-ø/ (Gottfried, 1984). While it has long been suggested that human perception of non-native sound contrasts is constrained by the phonological and phonetic influences from their native language (Best, 1995; Flege, 1995; Polka, 1991, 1992; Strange, 1995), little attention has been paid to the perception of non-native lexical tones (a suprasegmental feature) from the perspectives of the phonological and phonetic influences from native languages. The goal of this study is to investigate the issues in order to have a better understanding of the influence of native prosodic system (e.g., lexical tones) on adults' perception of non-native tones.

2. Background

2.1. Perception of tones

Lexical tones, similar to vowels, are perceived via normalization, an auditory skill that native listeners use to extract invariant tonal patterns and to ignore a speaker's absolute physical pitch values (i.e., fundamental frequency, F0) (Ching, 1984; Moore & Jongman, 1997; Wong & Diehl, 2003). As long as the perceived tonal patterns conform to those in their mental representations, native listeners will be able to perceive the tonal patterns and to identify the tones (Ching, 1984).

Among the perceptual cues of Mandarin tones - *F*0, duration and intensity, *F*0 cue is the primary perceptual cue for lexical tones (Gandour, 1994; Halle, Chang, & Best, 2004). In particular, both *F*0 height and *F*0 contour are the two essential dimensions of the *F*0 cue for the identification of tones by native Mandarin speakers (Gandour, 1984; Massaro, Cohen, & Tseng, 1985; Moore & Jongman 1997). In addition, the *F*0 turning point has been identified to be an important cue for distinguishing between Tone 2 and Tone 3 (Moore & Jongman, 1997). Besides *F*0 cues, durational and amplitude cues may exist as concomitants, but are not crucial to the

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after abstract only review

perception of Mandarin tones (Gandour, 1994). For example, previous studies (Ho, 1976; Howie, 1976) have suggested that Mandarin tones appear to have some intrinsic durational differences: Tone 3 tends to be the longest; and Tone 4 is the shortest. Tone 2 is generally shorter than Tone 3, but longer than Tone 1. However, Tseng (1990) found that duration was not a crucial acoustic parameter of Mandarin tonal production.

2.2. Perception of non-native tones

Non-native speakers of Mandarin perceive tonal categories different from native speakers (Halle et al., 2004; Leather, 1983), who can identify subtle differences between tones. Gandour (1983, 1984) identified that native English listeners tended to focus on the pitch height, while listeners from Chinese languages (e.g., Cantonese & Mandarin) focused more on pitch height and pitch direction when perceiving tones. Presumably, English has no lexical tones. It should also be mentioned that listeners with musical training backgrounds tend to perceive the lexical tones better than those who were not musically trained (Burnham & Brooker, 2002; Gottfried & Riester, 2000).

Studies on L2 tone perception suggest that linguistic experience of tones from native language assists listeners in perceiving non-native tones. Lee, Vakoch, & Wurm (1996) found that Cantonese listeners perceived Mandarin tones better than did English listeners. In their training study of a pair of Thai tones (mid vs. low), Wayland & Guion (2004) reported that naïve Mandarin Chinese listeners discriminated the Thai tones in the pretest better than did naïve English listeners. These studies suggested that their better perceptual performance was attributable to their greater linguistic experience in using pitch variations in distinguishing the lexical items. However, it is uncertain whether linguistic experience of tones always facilitates the perception of non-native tones, and whether the performance of listeners from different backgrounds will be different as a function of the use of tonality in their native languages.

In addition, Mandarin tones are not perceived and produced equally well by non-native listeners. It has been reported that some Mandarin tones are easier to be confused by other counterparts, because of the similarities in their tone contours, such as pitch onset and offset points, and contour shape (Gottfried & Suiter, 1997; Kiriloff, 1969; Miracle, 1989; Shen, 1989). Particularly, the pairs, Tone 2-Tone 3 and Tone 1-Tone 4, are frequently found to be difficult to discern for learners of Mandarin from non-tone languages (e.g., Australian, Dutch, and English languages). Since these findings are based on the performance of individual non-tone listener groups, whether the same patterns will also be observed from listeners of tone languages is still uncertain.

2.3. The Present Study

To address the above two issues, this study examined the perception of Mandarin tones (T1: high level [55]; T2: mid rising [35]; T3: falling rising [214]; and T4: high falling [51]) by two groups of naïve listeners from two tone languages, Hong Kong Cantonese and Japanese, which differ in degree of tonality use at the word level. Presumably, Cantonese speakers have greater experience in using pitch variations to distinguish lexical items than Japanese speakers. As described below, Cantonese has 6 phonemic tones and Japanese has only 2 in their systems. Also, in Cantonese, each syllable (i.e., word) is produced with a lexical tone, whereas not every Japanese monosyllable is produced with a pitch-accented pattern. More importantly, Cantonese listeners have extensive experience in using a wide range of variations of tones in Cantonese system, while Japanese listeners' experience in using pitch variations is limited to the combinations of the two tones, high (H) and low (L) (Yip, 2002). Below is a brief description of the two systems.

Cantonese, a typical lexical tone language, has six tones in the forms of seven pitch patterns (i.e., T1: high level/high falling [55/ 53]; T2: high rising [35]; T3: mid level [33]; T4: low falling [21]; T5: low rising [23]; and T6: low level [22]). Cantonese Tone 1 [55] has two allotones, a high falling tone [53] and a high level tone [55] (Bauer & Benedict 1997; Hashimoto 1972; Yip, 2002). The high falling tone is also well documented in a tone sandhi environment (Hashimoto, 1972; Bauer & Benedict, 1997). A high falling [53] becomes a high level [55] before a high falling or a high level. However, unlike Guangzhou Cantonese speakers, most Hong Kong Cantonese speakers "have lost the high falling tone, or use it in certain syntactic environments, or use it in free variation with high level" (Bauer & Benedict, 1997, p.167).

Japanese, a pitch-accent language, is a "subtype of tone language" with a few contrasting tones in its prosodic system (Yip 2002, p.4). For those two-mora words, the two tones (H and L) in its system can combine as HL or LH pitch-accented patterns that are similar to open syllable words of Chinese (both Mandarin and Cantonese), in term of syllable weight. (Note that: although it is well known that Tokyo dialect is different from Osaka dialect, but both dialects do use the HL and LH pitch accents for distinguishing lexical items, see McCawley, 1978; Takeki, 2004).

On the basis of the findings of previous studies, two hypotheses were tested. First, Cantonese listeners with greater linguistic experience in using pitch (variations) from their native language will perform better than Japanese listeners. Second, both Cantonese and Japanese listeners will also experience greater difficulties in perceiving Mandarin T1-T4 and T2-T3 pairs. If the confusions between the tones in the pairs are due to phonetic similarities of the target tones, similar difficulties should also be observed in the listeners of other tone languages.

3. The experiment

3.1. Participants

Two listener groups of 10 participants each were recruited: Hong Kong Cantonese (a tone language), and Japanese (a pitch-accent language). The Cantonese listeners (n=10) ranged in age from 18–26 years (M=21.7 years), and the Japanese listeners (n=10) ranged in age from 18-36 years (M=23.8 years). All native speakers of Cantonese and of Japanese were born and raised in their home countries, and came to Canada after the age of 15 years. They were all college and university studies at the time they took the experiment. All participants met two stringent criteria. They had neither prior Mandarin knowledge nor formal musical training. The goal was to examine to what extent phonological and phonetic influences from native prosodic systems affected the perception of non-native lexical tones.

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after abstract only review

3.2. Stimuli

There were 144 stimuli consisting of three Mandarin open syllables (*di, da*, and *du*) with four tones spoken by a female and a male speaker of Mandarin from Beijing (2 speakers x 3 syllables x 4 tones x 2 samples per tone x 3 repetitions). The intelligibility of the stimuli were evaluated by two additional native Mandarin speakers (22 and 27 years old, respectively), and all stimuli were identified by the raters.

3.3. Procedure

Before the test, all listeners were given a short familiarization section, in which they were given 24 tokens of Mandarin tones on other syllable words spoken by both speakers in this study. For the listening task, an identification paradigm was used to assess the listeners' perception of the four Mandarin tones. They were asked to respond the Mandarin tones from the options of four tones. The stimuli were presented in two block (1F; 1M) and presented to the listeners in a counterbalance order.

4. The results

The analysis was based on A-prime (*A*') scores and perceptual errors for Mandarin tones in six pairs, T1-T2, T1-T3, T1-T4, T2-T3, T2-T4, and T3-T4. (Note that: pre-analyses of the listeners' data did not find any significant gender effect between the speakers on the data, and therefore the gender factor was not considered in this study).

4.1. A-prime scores

The listeners' mean *A*' scores for the four Mandarin tones is presented in Figure 1. The listeners' *A*' scores ranged from 0.74 to 0.86.

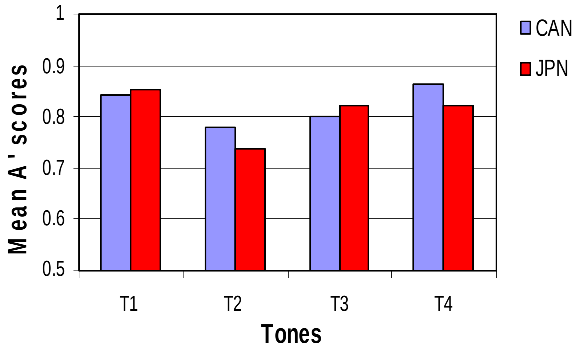

Figure 1: Mean *A*' scores by the listeners of the Cantonese and the Japanese groups.

The listeners' mean *A*' scores were submitted to a 2-way mixed-design ANOVA with L1 Group (CAN and JPN) as the between-subjects factor, and with Tone (T1, T2, T3, and T4) as the within-subjects factor. The results revealed only the effect of Tone was significant, $F(3,54)=6.196$, $p=.001$. Post Hoc LSD tests further revealed that the listeners' sensitivity to Tone 2 was significantly lower than did their sensitivity to other three tones ($ps < .05$).

4.2. Tonal errors

Listeners' mean tonal errors (confusions) are presented in Figure 2. Their mean errors were submitted to a 2-way mixed-design ANOVA with L1 Group as the between-subjects factor, and with Tone Pair (6 pairs: see Section 4) as the within-subjects factor. The results indicated that significant effect of Tone Pair was found, $F(5,90)=49.930$, $p<.001$. The interaction, L1 Group x Tone Pair, was also significant $F(5,90)=7.554$, $p<.001$; However, the L1 Group effect was not significant ($p>.05$).

To further explore the interaction, six individual t-tests were performed to examine the group difference for each tone pair. The results revealed that the Cantonese listeners made significantly more errors for the T1-T4 pair than did the Japanese listeners, $t(9)=3.639$, $p<.001$. Moreover, the error difference for the T2-T3 pair between the two groups was marginally significant, $t(9)=1.936$, $p=.08$. However, the Japanese made more errors for the T2-T4 pair than did the Cantonese listeners, $t(9)=-3.440$, $p<.001$.

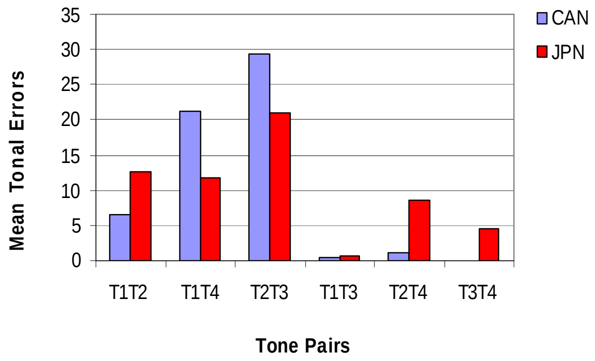

Figure 2: Mean errors for the six tone pairs by the listeners of the Cantonese and the Japanese groups.

In addition, two separate 1-way ANOVAs were performed for the Cantonese and the Japanese groups to investigate the effect of Tone Pair on the listener groups' mean errors. For the tonal confusions made by the Cantonese listeners, the analysis revealed that the Tone Pair effect was significant, $F(5,45)= 99.242$, $p<.001$. Post hoc LSD tests further revealed that the Cantonese listeners' mean tonal errors for the individual pairs, T1-T2, T1-T4 and T2-T3, were significantly different from their five counterparts, respectively ($ps < .05$). In addition, the error difference the T2-T4 and T3-T4 pairs was also significant ($p<.05$). These results also indicated that among the three pairs, there was a gradient pattern of mean errors among the three tone pairs: T2-T3 > T1-T4 > T1-T2.

For the Japanese listeners, the Tone Pair effect was also significant, $F(5,45)= 8.967$, $p<.001$. Post hoc LSD tests further revealed the following patterns for the Japanese listeners' mean tonal errors:

- T2-T3 > T1-T3, T2-T4, and T3-T4 ($ps<.05$);
- T1-T2 > T1-T3, T2-T4, and T3-T4 ($ps<.05$);
- T1-T4 > T1-T3 ($p<.05$);
- T1-T3 < T1-T2, T1-T4, T2-T3 and T2-T4 ($ps<.05$);

- T2-T4 < T1-T2, and T2-T3 pairs (*ps*<.05); but T2-T4 > T1-T3 (*p*<.05)
- T3-T4 < T1-T2 and T2-T3 (*ps*<.05).

5. Discussion

The results did not support the first hypothesis. Unlike the results of previous studies (Lee et al., 1996; Wayland & Guion, 2004), the performance of the Cantonese listeners was compatible with that of the Japanese listeners. This implies that the amount of linguistic experience in tonality use does not guarantee a facilitatory effect for the Cantonese listeners.

However, the results showed that both the listener groups, similar to those non-tone groups in previous studies (Gottfried & Suiter, 1997; Kiriloff, 1969), made considerable amount of errors for the T1-T4 and T2-T3 pairs. These findings support the second hypothesis. More interestingly, the results actually revealed some previously unnoticeable asymmetrical patterns of tone perception. Both the listener groups exhibited substantial confusion about the tones in the T1-T2, T1-T4, and T2-T3 pairs, but they made noticeably fewer errors for the other three pairs of tones, T1-T3, T2-T4, and T3-T4. It appeared that the former group of tone pairs involves tones sharing some degrees of similarity in terms of pitch contour patterns (e.g., pitch onset and offset and the pitch contours). For example, T1 and T2 both share the similar pitch height values for their offsets. In contrast, the second group of tone pairs involves tones showing lots of dissimilarities. For examples, the contour of T1 is totally different from that of T3: T1 has a high level pitch whereas T3 has a falling-rising pitch. These results imply that non-native tonal perception is under the influences of the phonetic characteristics of the target tones to a great extent.

However, there was a remaining unanswered question as to why the Cantonese listeners exhibited substantially more errors for the T1-T4 and T2-T3 pairs than did the Japanese listeners. The Perceptual Assimilation Model (PAM: Best, 1994, 1995) provides a theoretical basis for explaining the discrepancies in the listener groups' performance for tonal confusions.

The PAM posits that non-native speech perception is strongly affected by listeners' knowledge of native phonological systems. These naïve adult listeners have a strong tendency to perceptually *assimilate* non-native speech categories to the native categories that they perceive as the most similar (Best, 1994, 1995; Best, McRoberts & Goodell, 2001). However, listeners' abilities to discern the differences between non-native phones are predictable and dependent on how the non-native phones assimilate to native segments. The PAM provides a set of specific predictions of assimilations that incorporate both *phonological and phonetic influences* from the native languages. For example, when two non-native sound categories are perceived as a single native category, the discrimination between the two non-native categories will be very poor (*Single Category assimilation: SC*). In contrast, when the two non-native categories are assimilated to two individual native categories, the discriminations will be excellent (*Two Category assimilation: TC*). When both non-native phones assimilate to a single category, one tends to assimilate better than the other. Moderate to excellent discrimination is anticipated (*Category Goodness difference in assimilation: CG*). In the case that one non-native sound is assimilated as a native category while the uncategorized one falls in the phonetic space (outside any particular native categories), discrimination is expected to be good (*Uncategorized-Categorized assimilation: UC*).

In this study, the Cantonese listeners might have assimilated the perceived Mandarin tones to their Cantonese tones. In that case, Mandarin Tone 1 (high level [55]) and Tone 4 (high falling [51]) assimilated to the Cantonese Tone 1 (high level) only, because it has two allotones, high level [55] and high falling [53]. The consequence was that the Cantonese listeners frequently misidentified these Mandarin tones (target Tone 4 with Tone 1). This appeared to be the *SC* pair in the PAM. For the other case, Mandarin Tone 2 (mid rising [35]) and Tone 3 (falling rising [214]) might have been assimilated to Cantonese Tone 2 (high rising [35]). Possibility, the Mandarin Tone 2 assimilated to the Cantonese Tone 2 better than did the Mandarin Tone 3, and therefore the Cantonese listeners had a tendency to select Mandarin Tone 2 as the better match most of the time. There are two possible reasons for this tendency. First, the Cantonese high rising tone (Tone 2) is phonetically similar to Mandarin Tone 2 in terms of the *F*0 patterns (*F*0 height and shape). Both are described with the same tone letters [35] in the literature (Cantonese: Hashimoto, 1972; Yip, 2002, and Mandarin: Howie, 1976). The Cantonese Tone 2 is produced with a falling and a rising pattern (Bauer & Benedict, 1997; So, 1999), and so is the Mandarin Tone 2 (Fon & Chiang, 1999). Second, the Cantonese tonal system does not have a tone that is similar to Mandarin Tone 3 [214]. When a Cantonese speaker listens to a Mandarin Tone 3, the best candidate (the closest L1 tone) will be a Cantonese Tone 2, because Mandarin Tone 2 and Tone 3 share considerable similarities in their pitch contours, such as the dip and the rising portions. (Note that: there might have a possibility that the other candidate is Cantonese Tone 5. However, in terms of the degrees of the rising portion of Mandarin Tone 3, Cantonese Tone 2 appears to be the best candidate.) Thus, this could be the case of the CG pair. Therefore, according to the PAM, both the perception for the non-native contrasts in both SC and CG will be relatively more difficult for the perceivers.

As for the Japanese, T1-T4 and T2-T3 will be interpreted as two UC pairs. It is because Mandarin Tone 2 (mid rising) and Tone 4 (high falling) might have been assimilated to the Japanese LH and HL pitch-accented patterns, respectively. Mandarin Tone 1 and Tone 3 would be interpreted as uncategorized categories, because they did not assimilate to any tone or pitch-accented pattern in the Japanese prosodic system. According to the PAM, listeners' perceptions of uncategorized sounds are less influenced by their L1 systems, but it depends on how well those listeners perceive the similarities of the non-native contrasts. Perhaps, Tone 1 and Tone 3 have some phonetic properties (e.g., vowel duration and *F*0 patterns) that are relatively easy to perceive. Tone 1 involves high pitch with limited pitch movement; Tone 3 involves a low pitch value in the centre portion and is produced with longer vowel duration (e.g., Ho, 1976; Howie, 1976). Therefore, the pairs, T1-T4 and T2-T3, formed two UC pairs. According to the PAM prediction, listeners should be able to discriminate the non-native sounds of an UC pair quite well. The results of the present study were consistent with the prediction, because the Japanese listeners showed significantly fewer tonal errors for the tone pairs.

It should also be mentioned that the results showed that the Japanese listeners made significantly more errors for the

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006. Copyright, Australian Speech Science & Technology Association Inc.

Accepted after abstract only review

T2-T4 pair than did the Cantonese listeners. However, this may be due to the fact that the Cantonese listeners had relatively better discrimination between T2 and T4.

6. Conclusions

In examining the perception of Mandarin tones by two groups of tone languages (Cantonese and Japanese), which differed in the tonality use at the word level, it was found that both groups exhibited similar sensitivities to the Mandarin tones and tonal errors in general. Thus, linguistics experience does not provide the Cantonese listeners a facilitatory effect. In contrast, the results of the present study indeed indicated that the Cantonese listeners showed to have more difficulties in perceiving the tone contrasts between the tones in the T1-T4 and T2-T3 pairs than did the Japanese listeners. As discussed within the framework of the PAM, the Cantonese listeners' non-native tone perception appeared to be constrained by their native tone system. Specifically, the functional status of the tone contrast in the listeners' native phonology (*phonemic status*) or the similarity of some pitch contours (*phonetic characteristics*) of their existing L1 tones play a more significant role in perceiving of the non-native tones. Thus, this study contributes to the understanding of how the phonological and phonetic influences of adults' native prosodic systems affect their non-native tone perception. Moreover, the findings suggest that the similarity or dissimilarity of tonal contours (*phonetic characteristics*) also affect listeners' perception of tonal contrasts.

7. Acknowledgements

I would like to thank Catherine Best and the two reviewers for their comments and suggestions. All errors are mine. The work was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC).

8. References

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Berlin and New York: Mouton de Gruyter.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The Development of Speech Perception* (pp. 167-224). Cambridge, Massachusetts: The MIT Press.

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). 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. *The Journal of the Acoustical Society of America, 109*(2), 775-794.

Burnham, D., & Brooker, R. (2002). Absolute pitch and lexical tones: Tone perception by non-musician, musician, and absolute pitch non-tonal language speakers. In J. Hansen & B. Pellom (Eds.), *The 7th International Conference on Spoken Language Processing* (pp. 257-260). Denver , USA.

Ching, T. (1984). Lexical tone pattern learning in Cantonese children. *Language Learning and Communication, 3*(3), 317-334.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233-277). Timonium, Maryland: York Press.

Fon, J., & Chiang, W.-Y. (1999). What Does Chao Have to Say about Tones. *Journal of Chinese Linguistics, 27*(1), 15–37.

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

Gandour, J. T. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of. Chinese Linguistics, 12*, 235-261.

Gandour, J. T. (1994). Phonetics of tone. In R. E. Asher & J. M. Y. Simpson (Eds.), *The encyclopedia of language and linguistics* (pp. 3116-3123). Oxford: Pergamon Press.

Gottfried, T. L. (1984). Effects of Consonant context on the perception of French vowels. *Journal of Phonetics, 12*, 91-114.

Gottfried, T. L., & Suiter, T. L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics, 25*(2), 207-231.

Gottfried, T. L., & Riester, D. N. (2000). Preliminary results from study of pitch perception and Mandarin tone identification. *Journal of the Acoustical Society of America, 108*, 2604.

Hallé, P. A., Chang, Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*(3), 395-421.

Hashimoto, A. O.-K. Y. (1972). *Phonology of Cantonese.* Cambridge: Cambridge University Press.

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.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception & Psychophysics, 40*(4), 205-215.

Jamieson, D. G., & Morosan, D. E. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology, 43*(1), 88-96.

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

Leather, J. (1983). Speaker normalization in perception of lexical tones. *Journal of Phonetics, 11*, 373-382.

Lee, Y.-S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006. Copyright, Australian Speech Science & Technology Association Inc.

Accepted after abstract only review

comparison. *Journal of Psycholinguistic Research, 25*(5), 527-542.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of Acoustical Society of America, 89*(2), 874-886.

MacKain, K. S., Best, C., & Strange, W. (1981). Categorical perception of English /r/ and. /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2*, 369-390.

Massaro, D. W., Cohen, M. M., & Tseng, C.-y. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics, 13*(2), 267-290.

McCawley, J. D. (1978). What is a tone language? In V. A. Fromkin (Ed.), *Tone: A linguistics survey* (pp. 113-131). New York: Academic Press.

Miracle, C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of Chinese Language Teachers Association, 24*, 49-65.

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of Acoustical Society of America, 102*(3), 1864-1877.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *The Journal of Acoustical Society of America, 89*(6), 2961-2977.

Polka, L. (1992). Characterizing the influence of native language experience on adult speech perception. *Perception & Psychophysics, 52*(1), 37-52.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association, 24*(3), 27-47.

So, C. K. (1999). An acoustic analysis of Cantonese rising tones. *Journal of the Acoustical Society of America, 106*(4), 2153.

Strange, W. (Ed.). (1995). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium, Maryland: York Press.

Takeki, K. (2004). Tokyo and Osaka Japanese: is it possible to distinguish them by prosody alone? In *Proceedings of Workshop MIDL* (pp. 167-172). Paris, France.

Tseng, C.-y. (1990). *An acoustic phonetic study on tones in Mandarin Chinese*. Taipei: Institute of History & Philology Academia Sinica.

Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*(4), 681-712.

Werker, J. E., Gilbert, J. H. V., Humphrey, G. K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Developmental Psychology, 52*, 349-355.

Wong, P. C. M., & Diehl, R. L. (2003). Perceptual Normalization for inter- and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research, 46*, 413-421.

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

Proceedings of the 11th Australian International Conference on Speech Science & Technology, ed. Paul Warren & Catherine I. Watson. ISBN 0 9581946 2 9
University of Auckland, New Zealand. December 6-8, 2006.

Accepted after abstract only review