The Perception of Mandarin Tones by Thai and Indonesian Speakers

Raymond Wen-Chun Chow[1], Yi Liu[2], and Jing-Hong Ning[2]

[1]Hong Kong Community College, The Hong Kong Polytechnic University, HKSAR
[2]Chinese and Bilingual Department, The Hong Kong Polytechnic University, HKSAR
ccrchow@hkcc-polyu.edu.hk, yi.liu@polyu.edu.hk, jhning@polyu.edu.hk

Abstract

This study makes use of diverse acoustic features to comprehensively examine the effects of native language (L1) experience, tonal context, segmental context (consonant aspiration type and vowel height), and the intrinsic phonetic similarity on the perception of Mandarin tones by Thai and Indonesian speakers. Two perception tasks which are four-alternative forced-choice identification tests for stimuli presented in isolation and carrier sentences were conducted. Results showed that Thai listeners performed significantly better than Indonesian counterparts in both identification tests and the assertion of tonal language speakers having advantages over the non-tonal L1 speakers in acquiring a new tonal language was supported in this study. However, both groups share some similar error patterns which might be due to the intrinsic phonetic similarity between the target tones. The effect of segmental context appeared to be not significant, while the tonal context was found to exert contrary effect on Thai and Indonesian listeners.

Index Terms: prosodic background, Mandarin tone perception, acoustic features, segmental context, tonal context

1. Introduction

Previous cross-language research evidenced that listeners' language experience plays a crucial role in perception of L2 sounds at segmental level. It was well documented that learners' perceptual performance on non-native speech sounds was affected by their L1 phonetic and phonological properties [1, 2, 3, 4]. Recently, the studies of L1 influences on non-native (L2) speech sounds perception have been extended from segmental to suprasegmental level. Mandarin is regarded as one of the typical tonal languages in the world and its tonal system is always cited as the most difficult prosodic features to master for non-native speakers [5, 6, 7]. It has evoked a great deal of interest among linguists to investigate the difficulty in perceiving L2 tones [8, 9]. Some researchers attributed the failure of tone acquisition to the lack of tonal experience in learners' L1 [10, 11]. However, others found that the failure was due to the intrinsic phonetic similarity between target tone categories and there was no significant difference between the performance of tonal and non-tonal language speakers [12, 13, 14]. Beyond L1 background, segmental and tonal contexts have also been found to modulate the perception of tones [15, 16]. While most of the previous studies focused on examining how American, English, Cantonese, Japanese, Korean, and French perceived Mandarin tones, the perceptual performances among other languages have been seldom researched. It has not come up a robust conclusion whether tonal L1 will exert any positive, negative, or have no effects on Mandarin tone perception.

Among the available literatures which aimed to compare the nature of Mandarin tone perception between tonal L1 and non-tonal L1 speakers tended to select Cantonese speakers as their tonal L1 subjects. However, Cantonese and Mandarin share the same orthography and they are not good examples of mutually unintelligible languages, thus selection of Cantonese as subjects may render the findings unconvincing [13]. Selection of subjects outside the system of Chinese dialects can serve as a better sample for comparison. To deal with these gaps, this study selected two South Asian language speakers: Thai (tonal) and Indonesian (non-tonal) as subjects, aiming to comprehensively address whether tonal L1 speakers have advantage over the non-tonal counterparts on the perception of Mandarin tones. The interaction effects among consonant aspiration type, vowel height, and tonal context on the perception of Mandarin tones by tonal and non-tonal L1 speakers were also systematically compared.

1.1. Tonal patterns of Mandarin and Thai

Both Mandarin and Thai are tonal languages which utilize F0 height, pitch contour, and pitch slope as essential phonetic cues to discriminate lexical meaning. Mandarin has four lexical contrastive tones while Thai has five [17, 18, 19]. Although the tone letters of the five tones in Thai do not directly correspond to Mandarin, some considerable similarities in pitch contour can be identified between them. For example, similar to Mandarin T2 (35), Thai T4 (45) is a rising tone and they have an overlapping rising range (45 in tone letter values). Correspondingly, Thai T1 (33), T3 (41), and T5 (114 or 214) are phonologically similar to Mandarin T1 (55), T4 (51), and T3 (214) in terms of the pitch value and contour [20]. Given that Thai speakers can benefit from their L1 tonal system in the perception of Mandarin tones [21, 22], in this study, it is hypothesized that tonal L1 speakers (Thai) can outperform the non-tonal counterparts (Indonesian) in the perceptual performance of Mandarin tones and discrepancies on perceptual pattern would be discovered.

2. Method

2.1. Subjects

Altogether 18 Thai (9 male, 9 female), 18 Indonesian (8 male, 10 female), and 18 native speakers of Mandarin (8 male, 10 female) participated in this study. The Thai and Indonesian subjects were all year one undergraduate students who were studying the same Mandarin program at Jinan University, Guangzhou. The native Mandarin speakers from The Hong Kong Polytechnic University served as control group in this study. The average age of the three L1 groups was 20.4 years (SD=2.9) for Thai, 19.9 years (SD=2.3) for Indonesian, and 26.1 years (SD=3.4) for Mandarin group. Noted that at the moment the Thai and Indonesian students took part in this

study, they had come to China for one year. Although they were not naïve learners, they were far from being proficient in Mandarin and their exposure to Mandarin inputs was comparable. They had no or only limited music training and reported no speech or hearing impairments.

2.2. Stimuli

Four Mandarin consonant-vowel syllables, /ta/, /ti/, /tʰa/, /tʰi/ each carrying four Mandarin lexical tones, were selected as stimuli. No sonorant onset syllables were included in the experiments. Except T2 of /tʰa/, all syllables can form real words when combining with four lexical tones of Mandarin. Choosing the above four syllables to create experimental stimuli served two purposes. First, the consonants, vowels, and segmental sequences of these syllables can comply with Thai and Indonesian phonetic systems. This arrangement can minimize the interference from unfamiliar segments during the experiments. Second, since previous studies found that consonant aspiration type and vowel height can affect pitch perception [15, 16], using /t/, /tʰ/ and /a/, /i/ can enable us to investigate whether various segmental combinations would result in perceptual differences between the Thai and Indonesian listeners. Previous studies demonstrated that tone perception is modulated by speaker variability and speaker gender [23, 24, 25]. To avoid the practice effects, the stimuli were produced by two speakers (1 female and 1 male) and the tokens were presented to the subjects in different orders for the two identification tests.

2.3. Procedures

The perception task was comprised by two identification tests, including target tones presented in isolation and in the middle position of carrier sentences. For the isolated tones, a total of 64 tokens (4 syllables x 4 tones x 2 speakers x 2 repetitions) were yielded for each subject. The test was a four-alternative forced-choice identification task. As mentioned, the present study went beyond the perception of isolated tones by conducting another identification test which embedded the target tones in carrier sentences. Given that the carryover effect was greater than anticipatory for tonal coarticulation [26, 27], the four Mandarin tones were assigned as the preceding tones of the target syllable to generate all the four possible tonal contexts. The four tonal contexts were delivered by the carrier sentence "Wǒ shuō/xué/xiě/jì X zhè ge zì (I say/learn/write/remember the word X)" which X stands for the target syllable. According to the T3 sandhi rule in Mandarin, when a T3 syllable is followed by another T3, its pitch contour will change to a high-rising tone which is phonologically indistinguishable from a Mandarin T2 [28]. It is expected that the tonal context effect between the combinations of T2+T3 and T3+T3 are consistent, thus the T3+T3 combination was excluded from this experiment. In total, each subject was required to respond to 240 tokens (4 syllables x 4 tones x 4 preceding tones x 2 speakers x 2 repetitions – 16 T3+T3 combinations) in the identification test of the four Mandarin tones presented within tonal context. The subjects only received the audio sounds and no Chinese characters or Roman alphabet were provided during the tests.

3. Results

3.1. Perceptual performance for the target tones presented in isolation

Table 1 summarizes the tonal confusions of Thai and Indonesian listeners in the identification test of isolated tones.

Thai listeners got an overall accuracy rate of 91.7% (SD=17.9) in the test. Their accuracy rate for Mandarin T1, T2, T3, and T4 was 97.6% (SD=9.5), 80.6% (SD=24.5), 92% (SD=17.7), and 96.9% (SD=10.2) respectively. Friedman's ANOVA showed significant differences between the four tones (X^2 (3)=19.207, p<.0001) implying that the four Mandarin tones had elicited various degrees of difficulty to the Thai listeners. A 2-tailed Wilcoxon test was carried out for pair-wise comparisons within the L1 groups. The results indicated that T1 had a significantly higher accuracy rate than that of T2 (z=-4.159, p<.0001, Bonferroni adjusted) and T3 (z=-3.52, p<.0001, Bonferroni adjusted). No significant differences were found between the accuracy rate of T4 and the other three tones.

Table 1: Confusion matrices for the responses in the identification test of isolated tones.

	Response / Target	T1	T2	T3	T4
Thai Group	T1	97.60%	1.04%	0.35%	1.04%
	T2	1.74%	80.56%	16.32%	1.39%
	T3	0.35%	6.94%	92.01%	0.69%
	T4	0.35%	1.74%	1.04%	96.88%
Indon Group	T1	88.49%	7.57%	0.66%	3.29%
	T2	6.91%	59.54%	31.91%	1.64%
	T3	3.62%	5.26%	89.80%	1.32%
	T4	3.95%	5.59%	0%	90.46%

For the Indonesian subjects, the overall mean was 81.8% (SD=28.6). Their accuracy rate for Mandarin T1, T2, T3, and T4 was 88.5% (SD=23.3), 59.5% (SD=33.9), 89.8% (SD=18.3), and 90.5% (SD=23.8) respectively. The order of the accuracy rate T4>T3>T1>T2 (T4 is the highest) was different from the Thai subjects' order T1>T4>T3>T2. Friedman's ANOVA showed that the tone categories had significant effect on the accuracy rate of tone identification (X^2 (3)=3, p<.0001). Bonferroni adjustment revealed that the error rate of T2 was significantly higher than the other three tones: T1 (z=-6.043, p<.0001), T3 (z = -5.983, p<.0001), and T4 (z=-5.996, p<.0001).

Based on the overall performance, an individual Friedman's ANOVA was conducted to examine the effect of L1 group on the perceptual accuracy. The results showed significant effect of L1 group (X^2 (3)=110.260, p<.0001), implying that subjects' L1 background has a strong interaction towards the accuracy rate. Post hoc analysis with Wilcoxon signed-rank tests further indicated that the performance of Mandarin group was significantly better than Thai (z=-6.75, p<.0001, Bonferroni adjusted) and Indonesian groups (z=-8.936, p<.0001, Bonferroni adjusted) while Thai group's performance was significantly better than Indonesian group (z=-4.934b, p<.0001, Bonferroni adjusted). As shown in Figure 1, compared to the Thai group, Indonesian perceptual accuracy for T1 and T4 was much lower. Moreover, Indonesian subjects displayed a comparatively higher degree of confusion in the differentiation of T2, as they often misidentified target T2 as T3 (error rate: 31.91%). Wilcoxon signed-rank tests were conducted to further examine whether the three L1 groups encountered different levels of difficulty in distinguishing the four Mandarin tones. The statistical results showed that Mandarin performed significantly better than Indonesian group in identifying T1 (z=-3.695, p<.0001, Bonferroni adjusted), T2 (z=-6.355b, p<.0001, Bonferroni adjusted), T3 (z=-3.838, p<.0001, Bonferroni adjusted), and T4 (z=-3.241b, p<.05, Bonferroni adjusted). Additionally,

significant differences were found between Mandarin and Thai groups in the accuracy rates of T2 (z=-5.029, p<.0001, Bonferroni adjusted), T3 (z=-3.262, p<.05, Bonferroni adjusted), and T4 (z=-2.46, p<.05, Bonferroni adjusted). T2 appeared to be the most confusing tone for both Thai and Indonesian listeners. Meanwhile, compared to the Thai group, Indonesian subjects committed more errors in discerning T1 though the result failed to reach statistical significance.
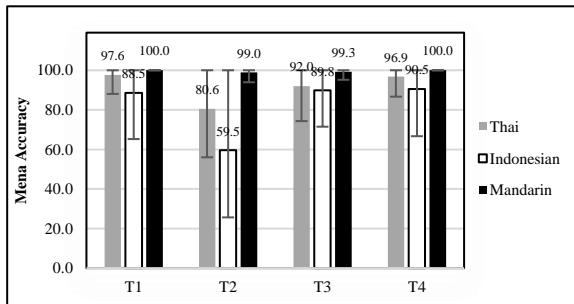

Figure 1: *Mean accuracy across the three L1 groups in the identification test of isolated tones.*

3.2. Perceptual performance for the target tones presented in context

For the target tones presented in context, the overall accuracy rate for Thai, Indonesian, and Mandarin subjects were 96.7% (SD=11.5), 76.1% (SD=32.6), and 99.1% (SD=6.1) respectively. Thai group had a nearly perfect performance in this test. They achieved 99.1% (SD=4.6) accuracy rate for T1, 92.9% (SD=18) for T2, 95.8% (SD=11.6) for T3, and 98.7% (SD=5.9) for T4. The hierarchy of tone accessibility from the least difficult to the most difficult tone was T1>T4>T3>T2 and this order was consistent with the finding for isolated tones. T1 was still the easiest while T4 appeared to be the most confusing tone to the Thai listeners. For the Indonesian listeners, their accuracy rate was 85.5% (SD=25.7) for T1, 65.5% (SD=32.4) for T2, 60.1% (SD=38.4) for T3, and 90.1% (SD=22.7) for T4. The accuracy rate order for the four Mandarin tones was T4>T1>T2>T3. This gradient was different from the finding in the identification test for isolated tones which T2 appeared to be the most difficult tone and T3 was the second difficult. This result implies that the tonal context might have exerted effects on the perception of Mandarin tones by Indonesian listeners.
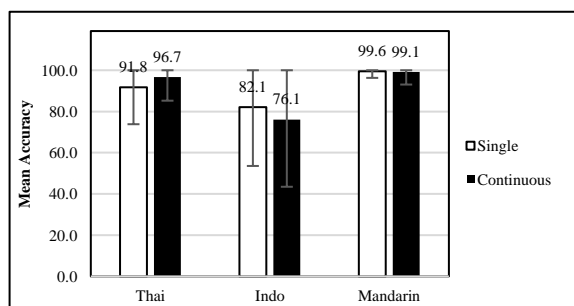

Figure 2: *Mean accuracy in the identification test for target tones presented in isolation and carrier sentences.*

Figure 2 compares the overall accuracy in the identification tests for target tones presented in isolation and carrier sentences. Wilcoxon signed-rank tests were carried out to examine the effect of tonal context. The results showed that Thai listeners performed significantly better in the identification test for target tones presented in context (z=-4.414, p<0.0001). However, the finding on the Indonesian listeners was inverse since their performance in the

identification test for target tones presented in context was significantly worse than the identification test for isolated tones (z=-2.136, p=0.033). No significant difference was found for Mandarin group.

Figure 3 depicts the mean accuracy of the four tonal contexts across the three L1 groups. Three individual Friedman ANOVAs were conducted to analyze the differences between the accuracy rates for the four tonal contexts within each L1 group. The results showed that there were no significant differences between the four tonal contexts for Thai (X² (3)=2.817, p=0.421) and Mandarin (X² (3)=1.286, p=0.733) groups, implying that the four tonal contexts exerted similar effects to Thai and Mandarin listeners. However, significant differences were found for Indonesian group (X² (3)=16.353, p<.005), indicating that the four tonal contexts had yielded different levels of difficulty to Indonesian listeners. Post hoc analysis further revealed that the accuracy rate for tonal context of T3 was significantly higher than T4 (z=-3.581, p<.0001, Bonferroni adjusted).
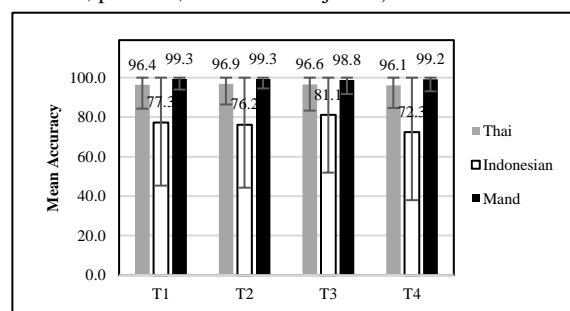

Figure 3: *Mean accuracy in the identification test of target tones presented within the four tonal contexts.*

For the effect of segmental context, Friedman ANOVAs were conducted to determine the influences of consonant aspiration on the perceptual accuracy within each L1 group. The results revealed that there were no significant effects of consonant aspiration (/t/ vs /tʰ/) among the Thai (z=-1.6, p=0.11), Indonesian (z=-1.2, p=0.23), and Mandarin (z-0.45, p=0.67) groups. Similar to the finding of consonant aspiration, no significant differences between the accuracy rate of low vowel /a/ and high vowel /i/ were identified among the Mandarin (z=-0.45, p=0.66), Thai (z=-0.18, p=0.88), and Indonesian (z=-1.5, p=0.13) groups.

4. Discussion

4.1. Effects of L1 on the perceptual performance

Mandarin and Thai are both tonal languages. In the light of their tonal inventories, Thai has a level tone (T1) and a high-falling tone (T3) which resemble to Mandarin T1 and T4. Although the tone letters of Thai T1 (33) and T3 (41) are different from Mandarin T1 (55) and T4 (51), their phonetic features are the closest among all the tone pairs. Thai T1 and T3 should be the best choice when determining the tonal assimilation of Mandarin T1 and T4 [20]. In this study, Thai listeners were shown to have almost perfect performance on discerning Mandarin T1 and T4 no matter the target tones were presented in isolation or carrier sentences. Previous studies also reported that Thai listeners could distinguish Mandarin T1 and T4 accurately in identification test [21, 22]. Nevertheless, among the literature concerning Mandarin tone acquisition, Mandarin T1-T4 tone pair was cited as perception problematic for Cantonese, English, Japanese, and Dutch speakers [8, 13, 14, 29, 30]. The asymmetric performances on

perceiving Mandarin T1-T4 contrast among different L1 groups imply that listeners' L1 prosodic background plays a significant role in non-native tone perception.

4.2. Effects of intrinsic phonetic similarity

According to the statistical results, the error rate of T2 was significantly higher than the other three tones for both Thai and Indonesian groups. The confusion matrices indicated that T2 and T3 were predominantly misidentified as each other among the confusion patterns. However, the underlying phonetic features of Mandarin T2 and T3 are comparable to Thai T4 and T5 (both Mandarin T2 (35) and Thai T4 (45) are high rising tones while the pitch features of Mandarin T3 (214) and Thai T5 (114 or 214) are low rising contour with an initial dipping portion) and it is predicted that Thai listeners should be able to discriminate the T2-T3 tone pair quite well. The discrepancies between the prediction and the findings of the current study suggest that the confusion of T2 and T3 might be L1-independent. That means the perception of T2 and T3 contrast is intrinsically harder than the other tone pairs regardless of the perceiver's L1. Consistently, previous studies on Mandarin tone perception also reported that the distinction between T2 and T3 appeared to be the most confusing for native speakers [7] and non-native speakers [9, 31]. The performance of Indonesian listeners' on identifying Mandarin T2 and T3 also supports the view that perception of T2-T3 contrast is interfered by the phonetic similarity between the target tones rather than listeners' L1 background. Similar to the finding on Thai group, Mandarin T2 was the most confusing tone category for Indonesian listeners and they frequently misidentified T2 as T3. Both groups demonstrated greater confusion for T2-T3 distinction, implying that this confusion pattern is irrespective of listeners' native languages. The failure of perception may be attributed to the intrinsic similarity between the phonetic features of Mandarin T2 and T3. Both Mandarin T2 and T3 start with a very close pitch onset and their pitch contours consistently exhibit an initial dip followed by a rising pattern at the final portion [19, 32]. The perceptual difficulty for Thai and Indonesian listeners may be due to these phonetic similarities.

Additionally, the similarity on tone duration may be another possible explanation for the T2-T3 confusion pattern. Mandarin T2 and T3 are not only similar in pitch height and contour but also in duration [19, 32]. Previous studies reported that the non-tonal L1 speakers rely more on syllable duration for Mandarin T2-T3 distinction [33]. In this study, the syllable duration seemed to exert similar effects on Indonesian subjects. According to the results of the two identification tests, although Indonesian listeners consistently made more errors in distinguishing T2 and T3, the accuracy rate for the identification of T3 presented within tonal context was significantly lower than in isolation because they frequently misidentified T3 as T2. The tonal context effect for T3 identification was found to be significant. In Mandarin, one of the T3 allophonic variations is that if a T3 syllable is placed at non-final positions, it will be pronounced as a low-falling tone (only the first half portion) and its tonal duration will be much shorter than an intact T3. Since the subjects of this study had not fully acquired the Mandarin tone system, it is possible that they were less sensitive to identify the subtle tonal contrasts, such as the degree of dipping pitch and the height of pitch offset, and they might utilize syllable duration as perceptual cue to help them to determine Mandarin tone categories. Given that Indonesian listeners made more T2 responses when T3 was presented in context (target tones with shorter duration) but more correct responses when T3 was presented in isolation

(target tones with longer duration), it is suggested that non-tonal L1 speakers depend more on duration cue when identifying T2-T3 contrasts.

4.3. Effects of segmental and tonal contexts

The current study hypothesized that the segmental context will affect listeners' perception of Mandarin tones. However, according to the statistical results, there was no significant difference between the accuracy rate for consonants /t/, /tʰ/ and vowels /a/, /i/. This finding suggests that subjects' perception of Mandarin tones was not affected by consonant aspiration and vowel height of the stimuli regardless listeners' L1 background. In other words, pitch value and contour still appeared to be the primary perceptual cues for tone identification in this study.

With regard to the effects of tonal context, only Thai listeners performed significantly better in perceiving target tones presented within tonal context than in isolation. In contrast, Indonesian group made significantly more errors when the target tones were embedded in carrier sentences. This asymmetric finding suggests that the tonal context can facilitate Thai listeners with their perception of tones, whereas it seemed to hinder the perception of tone by Indonesian listeners. This finding tends to support the view that tonal L1 speakers can benefit more from the tonal context when perceiving tones in connected speech. The statistical results further showed that there was no significant difference between the four tonal contexts for Thai and Mandarin groups. However, significant differences were found on Indonesian group, implying that the four tonal contexts would differ in their effects on tone perception by Indonesian listeners. Among the four tonal contexts, a preceding T3 seemed to facilitate the tone perception by Indonesian listeners whereas a preceding T4 might interfere with their tone identification. The reason for this perception pattern awaits further investigation.

5. Conclusion

In sum, there are four major findings in this study. First, the results of the identification tests showed that Thai listeners significantly outperformed the Indonesian counterparts in perceiving Mandarin tones, implying that the linguistic experience of lexical tones can facilitate non-native tone perception. However, the intrinsic similarities between the target tone categories and perceivers' L1 tonal pattern should also be considered when examining the effects of L1 prosodic background on non-native tone perception. Second, it was found that the tonal context exerted contrary effect to Thai and Indonesian listeners. Thai listeners performed significantly better when perceiving tones within tonal context than in isolation, suggesting that tonal context can assist tonal L1 speakers with their perception of Mandarin tones. Contrarily, the tonal context seemed to exert negative effects to Indonesian listeners. Third, it was found that the four tonal contexts differed in their effects on tone perception, but this finding confined to Indonesian listeners only. Fourth, the segmental contexts of consonant aspiration and vowel height were found to have no effect on Mandarin tone perception by Thai and Indonesian listeners. This finding supports the assertion that F0 height and contour serve as primary cues for Mandarin tone perception regardless the use of tonality in perceivers' L1.

6. References

[1] C. T. Best, "A direct realist view of cross-language speech perception," in *Speech perception and linguistic experience: Issues in cross-language research,* 1995, pp. 171-204. Timonium, Maryland: York Press.

[2] J. E. Fledge, "Second language speech learning: Theory, findings, and problems," in *Speech perception and linguistic experience: Issue in cross-language research*, 1995, pp. 233-277. Timonium, Maryland: York Press.

[3] L. Polka, "Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions," *The Journal of Acoustical Society of America,* vol. 89, no. 6, pp. 2961-2977, 1991.

[4] W. Strange (1995), "Cross-language studies of speech perception: A historical review," in *Speech perception and linguistic experience: Issue in cross-language research,* 1995, pp. 3-45. Timonium, Maryland: York Press.

[5] T. L. Gottfried and T. L. Suiter, "Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones," *Journal of Phonetics,* vol. 25, no. 2, pp. 207-231, 1997.

[6] C. Kiriloff, "On the auditory discrimination of tones in Mandarin," *Phonetica,* vol. 20, pp. 63-67, 1969.

[7] X. Shen and M. Lin, "A perceptual study of Mandarin tones 2 and 3," *Language and Speech,* vol. 34, pp. 145-146, 1991.

[8] X. X. Li, C. K. S. To, and M. L. Ng, "Effects of L1 tone on perception of L2 tone - a study of Mandarin tone learning by native Cantonese children," in *Bilingualism: Language and Cognition*, 2016, pp. 1-12. Cambridge: Cambridge University Press.

[9] Y. Wang, M. M. Spence, A. Jongman, and J. A. Sereno, "Training American listeners to perceive Mandarin tones," *The Journal of Acoustical Society of America,* vol. 106, no. 6, pp. 3649-3658, 1999.

[10] J. Liang and V. van Heuven, "Chinese tone and intonation perceived by L1 and L2 listeners," in *Tones and Tunes, Volume 2: Experimental studies in word and sentence prosody. Vol. 12-2. Phonology and Phonetics,* 2007, pp. 27-61. Berlin/NewYork: de Gruyter.

[11] R. Wayland and S. G. Guion, "Training English and Chinese listeners to perceive Thai tones: A preliminary report," *Language Learning,* vol. 54, no. 4, pp. 681-712, 2004.

[12] A. L. Francis, V. Ciocca, L. Ma, and K. Fenn, "Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers," *Journal of Phonetics,* vol. 36, no. 2, pp. 268-294, 2008.

[13] Y. C. Hao, "Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers," *Journal of Phonetics,* vol. 40, no. 2, pp. 269-279, 2012.

[14] C. K. So and C. T. Best, "Cross-language Perception of Non-native Tonal Contrasts: Effects of Native Phonological and Phonetic Influences," *Language and Speech,* vol. 53, no. 2, pp. 273-293, 2010.

[15] J. M. Hombert, "Perception of contour tones: An experimental investigation," *Proceedings of the 1st Annual Meeting of the Berkeley Linguistics Society.* pp. 221-232, 1975.

[16] Q. Zheng, "Effects of vowels on Mandarin tone categorical perception," *Acta Psychologica Sinica* vol. 46, no. 9, pp. 1223-1231, 2014.

[17] J. T. Gandour, "Perceptual dimensions of tone: Thai," in *Southeast Asian linguistic studies: Vol 3. Pacific Linguistics,* 1979, pp. 277-300. Canberra: Australia National University, Department of Linguistics.

[18] J. T. Gandour, "Tone perception in Far Eastern languages," *Journal of Phonetics,* vol. 11, pp. 149-175, 1983.

[19] M. Yip, *Tone.* Cambridge, UK: Cambridge University Press, 2002.

[20] T. Kwanrean, "The comparison and contrast of Mandarin and Thai tones," Thesis for Master's Degree, Yunnan Normal University, 2001.

[21] Y. Li, "English and Thai speakers' perception of Mandarin tones," *English Language Teaching,* vol. 9, no. 1, pp. 122-132, 2016.

[22] A. Rungruang and Y. Mu, "Mandarin Chinese tonal acquisition by Thai speakers," *Asian Social Science,* vol. 13, no. 5, pp. 107-115, 2017.

[23] Y.-H. S. Chang, Y. Yao, and B. H. Huang, "Effects of linguistic experience on the perception of high-variability non-native tones," *The Journal of Acoustical Society of America,* vol. 141, no. 2, pp. 120-126, 2017.

[24] C. Y. Lee, L. Tao, and Z. S. Bond, "Speaker variability and context in the identification of fragmented Mandarin tones by native and non-native listeners," *Journal of Phonetics,* vol. 37, pp. 1-15, 2009.

[25] G. Peng, C. C. Zhang, H.-Y. Zheng, J. W. Minett, and W. S.-Y. Wang, "The effect of intertalker variations on acoustic — perceptual mapping on Cantonese and Mandarin tone systems," *Journal of Speech, Language, and Hearing Research,* vol. 55, pp. 579-595, 2012.

[26] C. L. Shin, "The phonetics of Chinese tonal system," *Bell Laboratories,* Tech. Memor, 1987.

[27] Y. Xu, "Production and perception of coarticulated tones," *Journal of the Acoustical Society of America,* vol. 95, pp. 2240-2253, 1994.

[28] W. Wang and K. P. Li, "Tone 3 in Pekinese," *Journal of Speech and Hearing Research,* vol. 10, pp. 629-636, 1967.

[29] C. K. So, "Effects of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, English and Japanese," Doctoral dissertation, Simon Fraser University, 2006.

[30] J. Leather, "Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers," Paper presented at the *New Sounds: proceedings the 1990 Amsterdam Symposium on the Acquisition of Second-language Speech*, University Amsterdam, 1990.

[31] C. Shih, H. Y. Lu, L. Sun, J. T. Huang, and J. Packard, "An adaptive training program for tone acquisition," *Proceeding of Speech Prosody,* 2010.

[32] Y. R. Chao, *Mandarin primer*. Cambridge: Harvard University Press, 1948.

[33] Y. S. Chang, "Distinction between Mandarin Tones 2 and 3 for L1 and L2 listeners," *Proceeding of the NACCL,* vol. 1, pp. 84-96, 2011.