The Categorical Perception of Mandarin Tone 2 and Tone 3 by Tonal and Non-tonal Listeners

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

Previous studies on categorical perception (CP) of Mandarin tones focused on rising and falling continua. However, the influence of tonal experience on the categorical perception of Mandarin Tone 2 and Tone 3 remains unclear. By applying categorical perception paradigm, this study takes a further step to examine how L1 experience of Thai, Indonesian and Mandarin affect the CP along a Mandarin T2 and T3 continuum. The results showed that the three L1 groups differ significantly in the positions of the identification boundaries and boundary widths. Only Mandarin listeners exhibited categorical boundaries while Thai and Indonesian listeners demonstrated only psychophysical boundaries. However, both the identification and discrimination curves shown by the Thai listeners approximated the native speakers more closely than that of Indonesian counterparts. The results further pointed out that tonal and non-tonal L1 listeners process Mandarin tones differently while listeners' tonal L1 experience can facilitate the perception of non-native tone.

Keywords: categorical perception, linguistic boundary, psychophysical boundary, Mandarin tones

1. INTRODUCTION

1.1. Categorical perception

Mandarin is regarded as one of the typical tonal languages in the world and the challenge for mastering its tonal system is well documented in earlier literatures [8, 11, 21]. Some researchers attributed the failure of tone acquisition to the lack of tonal experience in learners' L1 [15, 28]. However, others found that the failure was due to the intrinsic phonetic similarity between target tone categories and the influence of L1 tone inventory [4, 10, 23]. Recently, the categorical perception (CP) has drawn increasing attention among the cross-language studies of Mandarin tones. In general, CP requires a co-occurrence of three perceptual characteristics: a sharp boundary between two categories; discrimination peaks at the category boundary; the observed discrimination performance can be predicted from the identification performance [18]. Like other segmental phonemes, tones must be phonologically categorized by listeners in order to achieve word recognition [5, 25]. By using rising and falling continua, previous studies found that tonal L1 listeners exhibited a higher degree of CP whereas non-tonal L1 listeners were able to process tone perception on the basis of their psychophysical factors only [2, 9, 17, 29].

1.2. Current study

Both Mandarin and Thai are tonal languages which utilize F0 height and pitch contour as phonetic cues to discriminate lexical meaning. Mandarin has four lexical contrastive tones while Thai has five [6, 30]. Although their tone inventories do not correspond to each other, some considerable similarities in pitch contour can be identified between them [12]. While most of the previous cross-language studies regarding CP of Mandarin tones focused on rising and falling continua [2, 9, 17, 29], the influence of tonal L1 experience, plus the different tone inventories on the CP of Mandarin T2-T3 remains unclear. To bridge the gaps, this study investigated how L1 experience of Thai (non-native tone speakers), Indonesian (non-tone speakers) and Mandarin (native speakers) affect the CP along a T2-T3 continuum. Given that Thai speakers can benefit from their L1 tonal system in the perception of Mandarin tones [14, 19], it is hypothesized that tonal L1 speakers (Thai) can perceive the continuum in a more categorical way, whereas the perception for non-tonal L1 speakers (Indonesian) might be more psychophysically based.

2. METHOD

2.1. Participants

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 participants were year one undergraduate students who had studied the same Mandarin program for 12 months at Jinan University, Guangdong. The native Mandarin participants served as control group in this study. The average age of the three L1 groups was 20.4 (SD=2.9) for Thai, 19.9 (SD=2.3) for Indonesian, and 26.1 (SD=3.4) for Mandarin. Noted that the non-native speakers were all naïve learners

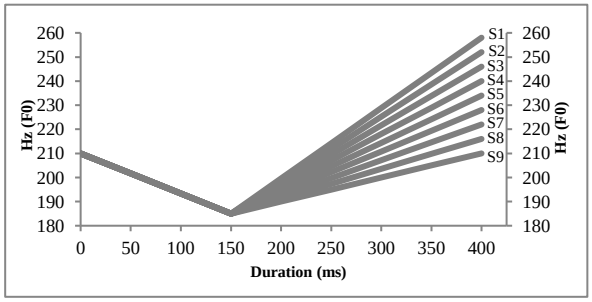
before they had come to China. Although they had studied Mandarin for one year, they were far from being proficient in Mandarin, thus enabling us to observe the effects from their L1 backgrounds. They had no or only limited music training and reported no speech or hearing impairments.

2.2. Stimuli

Mandarin syllable /ta/ was used to construct the T2-T3 continuum for this study. This syllable was chosen because its consonant, vowel, and segmental sequence comply with Thai and Indonesian phonetic systems so that can avoid the interference from unfamiliar segments. Wang and Li investigated how native speakers discriminate Mandarin T2 and T3 by using a T2-T3 continuum [26]. The results showed that under a fixed turning point condition, T2 response rate was significantly higher than T3 when the ending F0 was getting higher along the continuum. This study replicated similar paradigm to investigate the influence of L1 experience on CP of T2-T3 continuum.

Figure 1 shows a schematic diagram of the pitch contours of the 9 stimuli for the T2-T3 continuum. The stimuli were generated by the following procedures: 1) The Mandarin T2 and T3 of monosyllable /ta/ were first produced by a native female speaker in natural speech way. 2) The duration and the turning point of the target syllable were fixed to 400 ms and 150 ms respectively. 3) The starting, ending and turning pitch height of the middle stimulus (stimulus 5) were determined by the mean of the onset, offset and turning point F0 of the monosyllable /ta/ produced in T2 and T3. 4) Taking the stimulus 5 as anchor point, add and reduce the ending F0 by 6 pitch points to form other stimuli. The pitch-synchronous overlap and add (PSOLA) in Praat [1] was applied to resynthesize the stimuli.

Figure 1: Tone contours in the T2-T3 continuum.



2.3. Procedure

A two-alternative forced choice identification test and an AX discrimination test were conducted in order to obtain the essential characteristics of CP (boundary position, boundary width, and discrimination peak) for analysis. For the identification test, all the 9 stimuli along the /ta/ continuum were repeated 3 times, yielding 27 tokens for each participant. The stimuli were presented in random order and the participants were asked to identify whether the stimuli were Mandarin T2 or T3. For the discrimination test, a total of 23 pairs were presented in random order with a 400ms inter-stimulus interval (ISI). Among these pairs, 14 pairs consisted of two different stimuli separated by 2 steps on the T2-T3 continuum, in either forward (1-3, 2-4, 3-5, 4-6, etc.) or reverse order (3-1, 4-2, 5-3, 6-4, etc.), and 9 consisted of the stimuli of each step paired by itself (1-1, 2-2, 3-3, 4-4, etc). All pairs were presented twice, yielding 46 pairs in total. Participants were asked to judge whether the two target tones were the same or different. Instructions and practice trials were given before both tests.

3. RESULTS

Linear Mixed Model (LMM) was conducted to determine the effects of L1 experience on the CP parameters. When a main effect was significant, Wilcoxon signed-rank test (Bonferroni adjusted) was applied to make pair-wise means comparisons.

3.1. Results of boundary position

The CP position and width were assessed by Probit analyses of individual identification curves [3]. Table 1 summarizes the boundary position and width across the three L1 groups. LMM showed a significant main effect of L1 group ($F(1, 51)=18.73$, $p<0.0001$) indicating that the three L1 groups differed significantly in the boundary position. Pair-wise comparisons revealed that the boundary position was more toward to the stimuli with higher ending F0 for Indonesian group than for Mandarin ($z=-3.29$, $p<0.05$) and Thai group ($z=-3.55$, $p<0.0001$) while no significant difference was found between Mandarin and Thai group ($z=-1.02$, $p=0.31$).

Table 1: Derived position and width of categorical boundary for each L1 group.

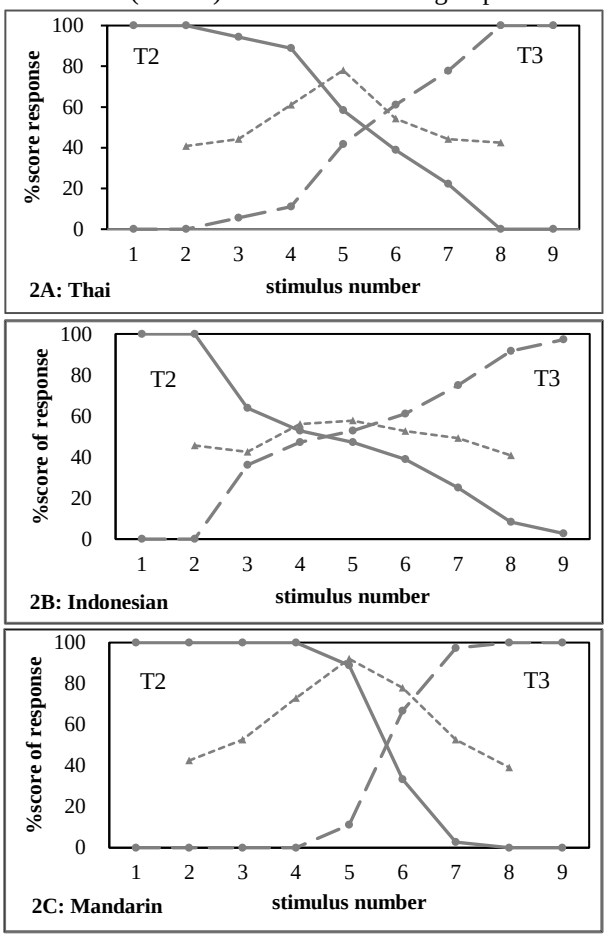
L1 group	Boundary postion (SD)	Width
Thai	5.35 (0.67)	1.62
Indonesian	4.35 (0.63)	2.56
Mandarin	5.63 (0.67)	1.23

3.2. Results of boundary width

The boundary widths across the three L1 groups were depicted in Figure 2 (A-C). A similar statistical analysis was conducted to examine the effects of L1 experience on the boundary width. The results showed that the boundary width differed significantly across the three L1 groups ($F(2, 51)=15.39$, $p<0.0001$). The boundary width for

Mandarin group was significantly narrower than Thai group ($z = -1.98$, $p < 0.05$) and Indonesian group ($z = -3.42$, $p < 0.05$) while Thai group exhibited a significantly narrower boundary width than Indonesian group ($z = -2.77$, $p < 0.05$).

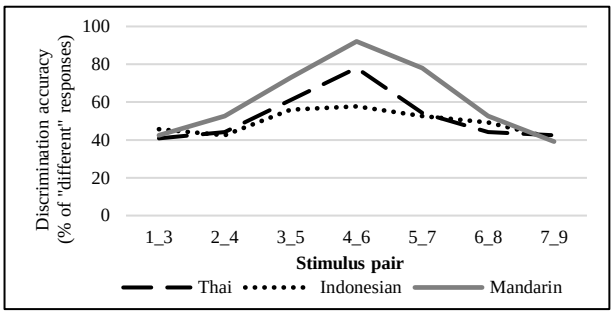
Figures 2 (A-C): Identification (solid) and discrimination (dashed) curves for each L1 group.



3.3. Results of discrimination test

Figure 3 depicts the discrimination curves pooled across L1 groups. The overall discrimination accuracy was 52.2%, 48.5%, 62.4% for Thai, Indonesian, and Mandarin group respectively. LMM indicated a significant main effect of L1 on the overall discrimination accuracy ($F(2, 51)=28.25$, $p < 0.0001$). Pair-wise comparisons revealed that the accuracy rate for the Mandarin group was significantly higher than that for the Thai group ($z=-3.29$, $p < 0.05$) and Indonesian group ($z=-3.7$, $p < 0.0001$) while no significant difference was found between Thai and Indonesian group. For the discrimination peak, the effect of L1 was not significant ($F(2, 51)=1.35$, $p = 0.27$), indicating that the discrimination peaks across the three L1 groups were in general the same (tone pair 4_6). However, significant difference in accuracy was found for this tone pair across the three L1 groups ($F(2, 51)=23.81$, $p < 0.0001$). Pair-wise comparisons indicated that Mandarin listeners significantly outperformed the Thai ($z=-2.72$, $p<0.05$) and Indonesian listeners ($z=-3.74$, $p<0.0001$). No significant difference was found between Thai and Indonesian group.

Figure 3: Two-step discrimination accuracy (%) pooled across the three L1 groups.



4. DISCUSSION

4.1. Position and width of category boundary

Peng et al. reported that non-tonal L1 listeners appeared to have broader boundary widths than tonal L1 counterparts for a rising and a falling pitch continuum whereas no significant difference was found on the boundary position [17]. These results are partially consistent with the current findings which show that the tonal and non-tonal L1 listeners differed significantly in both boundary position and width. We argue that this discrepancy is due to the different continuum types used in the experiments. Studies of T2-T3 perception found that native speakers rely much on the initial F0 fall and the timing of the turning point to distinguish the two tones [16, 21, 22, 26, 27]. However, in this study, the F0 of the tonal onset and turning point were fixed. This might render the Mandarin and Thai listeners unable to utilize the initial F0 fall as perceptual cue for T2-T3 distinction and they might need more steps to identify the stimuli as T3. In contrast, Indonesian listeners might be more sensitive to the pitch height than tone register and a slight fall on ending F0 might lead to a T3 response resulting in a significantly smaller boundary position. This finding further suggests that tonal L1 speakers pay more attention to the tone register and contour, but non-tonal speakers rely more on pitch height [6, 7].

Shen and Froud made use of a T2-T3 continuum to investigate CP of lexical tones by native English and Mandarin speakers. It was found that the English speakers required more steps (lower F0 for the turning point and end point) to make a T3 response [20]. This result is contrary to our finding that the boundary position for Indonesian listeners (non-tonal) is more toward to T2 Tonal Variants. This asymmetrical finding may provide some insights to the influence of listeners' L1 tone inventory. Although the low falling-rising tone (213)

in Thai is phonologically similar to Mandarin T3, they differ in the pitch register at the ending part of the tone. Such difference in tonal registration might render the Thai listeners to require more steps (lower ending F0) to identify the Tonal Variants as T3 and this might address why the boundary position for Thai group was found to be significantly larger than Indonesian counterparts. This finding further suggests that CP of tone is shaped by the tonal system of listeners' L1 [2, 9, 17, 29].

For the boundary width, as shown in Figure 2 (A-C), Mandarin group's category boundary is the sharpest and a clear cut can be seen between Tonal Variants 5 and 6, implying that their perception for the T2-T3 continuum was highly categorical. Whereas no sharp categorical boundaries were found on Thai and Indonesian groups, indicating that their perceptions were rather psychophysically based. A considerable amount of literature found that Mandarin T2 and T3 were predominantly misidentified as each other. Most studies attributed this high confusability to the similar F0 height and contour between the two tones [10, 13, 23]. In this study, the Thai and Indonesian listeners had come to China for only one year and it is possible that they were less sensitive to the subtle tonal contrasts, such as the degree of dipping pitch, turning point position, and height of pitch offset. This speculation is supported by the finding of the psychophysical boundary for the two groups.

4.2. Discrimination curve

In this study, although the overall discrimination accuracy for Mandarin group is significantly higher than Thai and Indonesian group, the actual value (62.4%) is not as high as expected. As shown in Fig. 3, the discrimination curve reaches the peak at the category boundary (between the S4 and S6) while it is relatively flat at both ends of the curve. The relatively low discrimination accuracy found on Mandarin group can be explained by the CP hypothesis that native speakers were found to be less sensitive within a category but clearly perceive differences between categories [24].

Fig. 3 shows that the discrimination curves for Mandarin and Thai groups are clearly bell shaped with a clear discrimination peak at tone pair 4_6. However, for Indonesian group, the curve is rather flat and only two fuzzy peaks are located at tone pairs 3_5 and 4_6. CP requires the discrimination peaks to be well aligned with identification boundaries and the discrimination could be predicted by the identification performance [25, 29]. As shown in 2(A-C), the discrimination peak for Mandarin and Thai listeners is well aligned with categorical boundaries whereas Indonesian group demonstrated similar accuracy rate for each tone pair (around 60-70%) and no clear peak was observed near the category boundary suggesting that they were responding to the psychophysical difference between Mandarin T2 and T3. This result could also be predicted by the broader boundary width exhibited by Indonesian group.

Although Thai and Indonesian listeners exhibited psychophysical nature on the perception of Mandarin T2-T3, their results differed significantly in boundary position and boundary width. Compared with Indonesian group, the identification curves for Thai group change more abruptly from one category to another and Thai's discrimination peak is aligned with corresponding category boundary. As the CP results found on the Thai listeners approximated the Mandarin listeners (native) more closely, it seems that Thai's perception for Mandarin T2 and T3 was more robust than Indonesian counterparts. We speculated that this asymmetrical result is due to the effects of L1 experience. In Thai, the high-rising (45) and low falling-rising tone (213) are similar to Mandarin T2 (35) and T3 (214) in terms of F0 and pitch contour. This intrinsic similarity between the tone inventories might facilitate Thai listeners' CP of Mandarin T2-T3. This finding is consistent with several studies that tonal L1 speakers perceive tones more categorically but the CP for non-tonal speakers is rather psychophysically motivated [2, 9, 17, 29].

5. CONCLUSION

This study aims at investigating whether tonal L1 speakers can perceive Mandarin T2-T3 in a more categorical way than non-tonal counterparts. Given that the CP pattern of Thai listeners approximated the native speakers more closely, this study suggests that Thai listeners can perceive Mandarin T2-T3 contrast more categorically and their tonal L1 experience can facilitate the development of their sensitivity to Mandarin T2 and T3. There might be a gradient difference of categoricity between the Thai and Indonesian listeners.

The stimuli used in this study differed in the ending F0 only. However, not only the ending F0 but also the F0 difference between tonal onset and turning point, the duration between tonal onset and turning point would have effects on T2 and T3 perception [16, 21, 22, 26, 27]. It will be more ideal if the present study can apply synthetic stimuli with the stimuli duration, onset and ending F0, the degree of the initial fall, and the turning point position of pitch contour systematically manipulated so that can ensure a more comprehensive and robust comparison between the CP nature of tonal and non-tonal L1 listeners. These parameters should be taken into account in designing future research.

6. REFERENCES

[1] Boersma, P., Weenink, D., 2009. Praat: Doing phonetics by computer. <http://www.fon.hum.uva.nl/praat/>.

[2] Chang, Y. C., Halle, P. A., Best, C. T., Abramson, A. 2008. Do non-native language listeners perceive Mandarin tone continua categorically? *Proc. 8th Phonetics Conference of China and the International Symposium on Phonetic Frontiers*, Beijing.

[3] Finney, D. J. 1971. Probit analysis. Cambridge, UK: Cambridge University Press.

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

[5] Francis, A. L., Ciocca, V., Ng, B. K. 2003. On the (non)categorical perception of lexical tones. *Perception & Psychophysic.* 65(7), 1029-1044.

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

[7] Gandour, J. T., Harshman, R. A. 1978. Cross-language differences in tone perception: A multidimensional scaling investigation. *Language and Speech*. 21, 1-33.

[8] 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), 175-184.

[9] Hallé, P. A., Chang, Y. C., Best, C. T. 2004. Identification and discrimination of Mandarin Chinese by Mandarin Chinese vs. French listeners. *Journal of Phonetics.* 32, 395-421.

[10] Hao, Y. C. 2012. Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*. 40(2), 269-279.

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

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

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

[14] Li, Y. 2016. English and Thai speakers' perception of Mandarin tones. *English Language Teaching.* 9(1), 122-132.

[15] Liang, J., van Heuven, V. 2007. 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*. Berlin/NewYork: de Gruyter, 27-61.

[16] Moore, C. B., Jongman, A. 1997. Speaker normalization in the perception of Mandarin Chinese tones. *Journal of Acoustical Society of America.* 102, 1864-1877.

[17] Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., Wang, W. S. Y. 2010. The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics.* 38, 616-624.

[18] Repp, B. H. 1984. Categorical perception: Issues, methods, findings. In: Lass, N. J. (eds), *Speech and language: Advances in basic research and practice (Vol. 10)*. New York: Academic Press, 243-335.

[19] Rungruang, A., Mu, Y. 2017. Mandarin Chinese tonal acquisition by Thai speakers. *Asian Social Science.* 13(5), 107-115.

[20] Shen, G., Froud, K. 2016. Categorical perception of lexical tones by English learners of Mandarin Chinese. *Journal of Acoustical Society of America.* 140, 4396-4403.

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

[22] Shen, X., Lin, M., Yan, J. 1993. F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of Acoustical Society of America.* 93, 2241.

[23] So, C. K., Best, C. T. 2010. Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech.* 53(2), 273-293.

[24] Stagray, J., Downs, D. 1993. Differential sensitivity for frequency among speakers of a tone and nontone language. *Journal of Chinese linguistics.* 21, 143-163.

[25] Wang, W. S. Y. 1976. Language change. *Annals of the New York Academy of Sciences*. 208, 61-72.

[26] Wang, Y. J., Li, M. L. 2010. The effects of tone pattern and register in perceptions of Tone 2 and Tone 3 in Mandarin. *Xinli Xuebao Acta Psychophysica*. 9, 899-908.

[27] Wang, Y. J., Qin, X. H. 2015. The identification and discrimination of Tone 2 and Tone 3 in Mandarin: The influence of experimental design on categorical perception of tones. *Yuyan Kexue.* 14(4). 337-352.

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

[29] Xu, Y., Gandour, J. T., Francis, A. L. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of Acoustical Society of America.* 120(2), 213-220.

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