A Categorical Perceptual Study on Mandarin Tones by Zaiwa language speakers

Yao Lu
Renmin university of China libraries; Department of Chinese language and literature, Peking University
luyiru2012@163.com
Jiangping Kong
Department of Chinese language and literature, Peking University; jpkong@pku.edu.cn

ABSTRACT

This paper investigated how the Zaiwa language speakers perceive Mandarin tones. Therefore, identification and discrimination experiments were conducted of the four Mandarin tone continua with Zaiwa Language speakers. The results suggest that Zaiwa speakers cannot establish a categorical perception of T2-T3 and T3-T4(T4) continua, but for other continua is categorical. Further analysis indicates that phonation information and F0 of the tone system in Zaiwa language both have a significant influence on the perception of Mandarin tones.

Keywords: Mandarin tones, categorical perception, Zaiwa language, phonation, language experience

1. INTRODUCTION

Categorical perception (CP) has been one of the most extensively studied phenomena in speech perception for many years because it is believed to reflect some fundamental aspects of how speech is processed [1]. Since 1970s, more and more attention has been put to the suprasegmental features of speech, especially the tone. Mandarin has four lexical tones which can be described phonetically as a high level tone (T1), a high rising tone (T2), a low-falling-rising tone (T3) and a high falling tone (T4). The value of the four tones is described as 55, 35, 214 and 51 on Chao's five-point scale [2], with 1 indicating the lowest F0 level and 5 the highest. Though not without controversy, the consensus opinion has been that the continua between T1-T2, T1-T3, T2-T4 and T1-T4 are perceived categorically by native speakers, while the continuum between T2-T3 is not, as it is difficult to find an accuracy peak at the category boundary [3,4,5,6].

Similar CP research has been conducted on second language (L2) tone perception. According to the perceptual assimilation model (PAM) [7] and speech learning model (SLM) [8, 9], different inventories of tones influence L2 tone perception, and it largely depends on differences between the L1 and L2 tone systems [10,11,16]. However, there has been a lack of study on whether and how different tone inventories affect tone perception in the framework of CP. There are many tonal languages and dialects used within China, and the tone systems and phonation types of some minority languages are even more complex than Mandarin. Therefore, how do tonal language speakers perceive Mandarin in the context of Mandarin popularization?

Thus, in this paper we chose to study Zaiwa speakers in the Dehong Autonomous Prefecture in regard to their perception of Mandarin tones and the influence of their mother tongue in the framework of CP. (Zaiwa – also known as Atsi – is a Burmese language spoken in parts of China and Burma.) Zaiwa's tonal system is unique for its phonation information, which is matched with tones to distinguish lexical meanings together. The previous acoustic experiment revealed that Zaiwa language contains six monosyllabic tones – two level tones, one rising tone and three falling tones – and their values are 55, 44, 35, 51, 31 and 21. The high rising tone (35) must be combined with vowels of modal voice, while other tones can be combined with vowels either of pressed and modal voice [12].

Regarding to the F0 aspect, high falling (51), high level (55) and high rising (35) tones exist in both Mandarin and Zaiwa language. In contrast, the Mandarin low-falling-rising tone (214) and the Zaiwa low falling tone (31&21) appear to lack acoustic equivalents in the other language. Regarding to the phonation aspect, Although the phonation information of Mandarin tones does not distinguish lexical meanings, it also changes with the tones. Kong [13] indicated that the most notable phonation characteristic was the irregular glottal periods (creaky voice) found at the middle of the low-falling-rising tone, which is similar to the vowels of pressed voice in the Zaiwa language. The previous study found that both F0 and phonation contribute to the native speaker's perception of Zaiwa tones, but to varying degrees in different environments, and the lack of phonation will lead to the "F0 neutral perception" phenomenon [12].

According to the comparison of the two tonal systems and the previous study, we assumed that the Zaiwa speakers may have difficulty in establishing a CP of Mandarin's low-falling-rising tone and high rising tone. Moreover, the phonation type in Zaiwa may also have a significant influence on their perception of Mandarin tones.

2. EXPERIMENT

2.1. Subjects

Twenty-eight native Zaiwa language speakers (ZL) participated in our experiment. The participants – 14 men and 14 women – were all from Huyu village, Ruili County, Dehong Autonomous Prefecture. Their ages range from 18-48, and at least received a primary education.

2.2. Stimuli

The stimuli were derived from the Mandarin syllable /da/ with four Mandarin tones (T1: da1, "take"; T2: da2, "reach"; T3: da3, "hit"; T4: da4, "big") and pronounced by a female native speaker. Four naturally produced /da/ syllables were normalized to a duration of 500 ms. Using Praat software, tone continua were manipulated from each of the four /da/ syllables to the other one. Take T1-T2 for example, 11 pitch points of F0 were extracted from "da1" and "da2" to plot their pitch contour. Then nine pitch contours were computed between the two syllables' original F0 contour by dragging the 11 pitch points equidistantly. Since the phonation cues also have an influence on CP, "da1" and "da2," were used respectively as the original synthetic sample. Therefore, there were two continuum subgroups with the same change of F0 parameters but different phonation information in each group. Finally, 11 steps of stimuli in each of the 12 pairs of tone continuum were synthesized into new sound files in Praat. The pitch range of T1-T2 is from 172.72 Hz to 297.58 Hz, T1-T3 is from 126.65 Hz to 294.91 Hz, T1-T4 is from 102.58 Hz to 316.92 Hz, T2-T3 is from 111.99 Hz to 297.58 Hz, T2-T4 is from 122.08 Hz to 302.5 Hz, T3-T4 is from 113.49 Hz to 303.89 Hz。

2.3. Procedure

In the identification task, participants listened to stimuli from the 12 continua in random order. The 11 stimuli of each continuum were repeated two times in the whole identification session, forming a total of 132 trials. During each trial, there were two Chinese characters on the screen of the computer, and participants were instructed to choose which character they heard. In the discrimination task, stimuli were presented in pairs with a 500-ms inter stimulus interval. A total of 324 pairs were presented in random order. Of these pairs, 216 consisted of two different stimuli separated by two steps. The remaining 108 pairs contained one of the middle nine stimuli of the continuum paired with itself. After hearing each pair, participants were instructed to judge whether the stimuli were the same or different.

2.4. Data analysis

We obtained results from each subject based on the two essential characteristics of CP: 1) sharp category boundary; 2) corresponding discrimination peak to the category boundary [1,4,5,14,15,16]. The identification score was defined as the percentage of responses with which participants identified that stimulus as being either 'Sound1' or 'Sound2'. Then the boundary position and boundary sharpness were measured by applying logistic regression analysis as Xu et al. [1]. To obtain the discrimination scores for each pair, we also used the formula described by Xu et al. [1] and then Repeated Measure ANOVA was applied as Yu [5] to determine whether the discrimination curve of each group has a significant peak. We used Excel 2010, IBM SPSS Statistics 20.0 and The R Project for Statistical Computing for our data analysis.

3. RESULTS

The category boundary position and width obtained by probit analysis of the 12 groups were presented in Table 1. Smaller width indicates sharper category boundary [1,4,5,14,15,16]. Repeated Measure ANOVA showed that results of the subgroups based on the same change of F0 parameters but different phonation information in T2-T3, T2-T4 and T3-T4 groups were significant differently($p<0.05$), whereas in other groups were not ($p>0.05$). Therefore, we merged the data of each two subgroups in T1-T2, T1-T3 and T1-T4 together. Moreover, the boundary width of the subgroups in T2-T3 varies significantly($p<0.001$). Besides, Repeated Measure ANOVA also showed that significant discrimination peak corresponding to the identification boundary can be found in the perception results of T1-T2, T1-T3, T1-T4, T2-T4 and T3-T4(T3) continua, but not in the T2-T3 (T2), T2-T3 (T3) and T3-T4(T4) continua.

Table 1: Categorical boundary position, width of identification task and discrimination peak

Group	Position	Width	Peak
T1-T2	5.01	2.92	significant
T1-T3	5.11	2.82	significant
T1-T4	4.36	1.93	significant
T2-T3(T2)	6.96	2.74	none
T2-T3(T3)	6.28	14.83	none
T2-T4(T2)	6.36	1.67	significant
T2-T4(T4)	6.84	1.67	significant
T3-T4(T3)	5.57	1.93	significant
T3-T4(T4)	6.72	1.91	None

4. DISCUSSION

4.1. Zaiwa speakers' perception modes of Mandarin tones.

According to the figures and data, it can be said that the identification curve is very sharp with a boundary width between 0-2 and relatively sharp between 2-4. These two types of identification results can both indicate a clear identification of the tone category of each tone continuum at the linguistic level in this study. However, a boundary width larger than 4 indicates it was difficult for the participants to discriminate the tone continua into two categories. Therefore, the data indicates that the Zaiwa speakers in our study could not distinguish the T2-T3 (T3) continua. For the assessment of the discrimination peak, discrimination peaks are not found in the discrimination curves of T2-T3 (T2), T2-T3 (T3) and T3-T4(T4) continua by the statistical method, but can be found in other continua, although the peak types and significance are slightly different. In general, at the behavioural level, the Zaiwa speakers perceived T2-T3 (T2), T2-T3(T3), T3-T4(T4) continua continuously and other tone continua categorically.

4.2. In terms of CP, how does the Zaiwa tone system affect tone perception of Mandarin?

The outcomes of this study align with previous findings and our hypothesis that Zaiwa speakers do have difficulty in establishing CP of T2-T3 continuum because of their similarity in some conditions. Interestingly, results of the repeated measures ANOVA showed that continuum subgroups based on the same change of F0 parameters – but with different phonation information in the T2-T3 continuum – were perceived significantly differently($p<0.001$). As we already controlled all the variables and only changed the F0 when synthetizing the stimuli, so the result of this differences also align with previous findings and our hypothesis that the phonation type in Zaiwa do have a significant influence on their perception of Mandarin tones. As reviewed earlier, the creaky voice in Mandarin T3 is similar to the pressed voice of vowels in Zaiwa language's high level tones (55 & 44). In addition, there is only one rising tone (35) in their perceived space and must be combined with vowels of modal voice. Thus, when the original synthesized sample was Mandarin T2 with modal voice, the Zaiwa speakers could clearly identify it as its F0 and phonation information are both similar to Zaiwa high rising tone (35), and it was discriminated from Mandarin T3, which lacks a counterpart in the Zaiwa tonal system clearly. Conversely, if the continua were synthesized by Mandarin T3 with pressed voice-although the tone changed from 214 to 35-the phonation information was still pressed, and this made it difficult for the Zaiwa speakers to identify a pressed high rising tone from a low-falling-rising tone. Findings are consistent with the perception study of Zaiwa tone perception [12], which suggest that the phonation and F0 should be both considered crucial cues for tone perception in terms of CP.

5. CONCLUSIONS

The current study investigated how Zaiwa speakers perceived the four Mandarin tones in terms of CP by conducting identification and discrimination experiments. The results suggest that the Zaiwa speakers could not establish typical CP of T2-T3 and T3-T4(T4) continua, but for other continua is categorical. Further analysis indicated that phonation information and the F0 of the tone system in Zaiwa both have a significant influence on the perception of Mandarin tones.

6. ACKNOWLEGEMENT

The work described in this paper was supported by grants from Ministry of Education of China. (Project Name: Language ontology research based on multi-modal; Project Number: 17JJD740001)

7. REFERENCES

[1] 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 the Acoustical Society of America. 120(2):1063-1074.

[2] Chao, Y.R.1948.Mandarin primer: An intensive course in spoken Chinese. Cambridge, MA: Harvard University Press

[3] Wang, W. S-Y. 1976. Language change. Annals of New York Academy of Sciences. 280(1):61-72.

[4] Yang R X. 2015.The role of phonation cues in Mandarin tonal perception [J]. Journal of Chinese Linguistics, 43(1):453-472.

[5] Yu Q. 2017. Dialectal background and Mandarin tone perception [D]. Peking University. PhD dissertation

[6] Qin X H. 2012.The influence of language experience on categorical perception of Mandarin tones-based on the research of Mandarin and Cantonese [D]. Peking University. Master's dissertation

[7] Best C T. 1995.A Direct Realist View of Cross-Language Speech Perception [J]. Speech perception and linguistic experience: Issues in cross-language research, 171-204.

[8] Flege, J. E. 1992. Speech Learning in a Second Language [A]. Edited by Ferguson, C., Menn, L. &Stoel-Ga

mmon, C. Phonological Development: Models, Research, and Implications[C]. Timonium, MD: York Press

[9] Flege, J.E.1999.The Relation between L2 Production and Perception[R]. In The Proceedings of the 14thICPhS (San Francisco)

[10] 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.

[11] Wu X, Munro M J, Wang Y. 2014.Tone assimilation by Mandarin and Thai listeners with and without L2 experience [J]. Journal of Phonetics, 46(1):86-100.

[12] Lu Y, Kong J P. Acoustic and Perceptual Studies on the Zaiwa Tones[J], Minor languages of China.2019,1:55-65.

[13] Kong J P.2001. On Language Phonation [M]. Central University for Nationalities Press.

[14] Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. 1957. The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology. 54(5):358-368.

[15] Studdert-Kennedy, M., Liberman, A. M., Harris, K. S., & Cooper, F. S. 1970. Motor theory of speech perception: a reply to Lane's critical review. Psychological Review. 77(3):234-249.

[16] 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(4):616-624.

[17] Jiang C, Hamm J P, Lim V K, et al. Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics[J]. Memory & Cognition, 2012, 40(7):1109-1121.