A Perceptual Study on the Distinctive Features of Entering Tones in Guangzhou Cantonese

Dong Han[1], Mingyang Yu[2]

[1]*Hong Kong Shue Yan University (Hong Kong),* [2]*City University of Hong Kong (Hong Kong)*
dhan@hksyu.edu, mingyanyu5-c@my.cityu.edu.hk

Background: Acoustic experiments indicate that in Cantonese the monosyllables of the high-level tones (HLT) and high-level entering tones (HLET) share pitch values, but differ in duration, as do the mid-level tones (MLT) and mid-level entering tones (MLET), and the low-level tones (LLT) and low-level entering tones (LLET) [1]. However, various studies have reported differences in the contribution of the features of the stop coda and the short duration to the perception of entering tones in Guangzhou Cantonese [2,3,4]. To better investigate the role of the two features in the perception of entering tones, the present study conducted three perceptual experiments on 12 native Guangzhou Cantonese speakers (6 male, 6 female) by manipulating the duration of three pairs: namely, "HLT - HLET", "MLT - MLET", and "LLT - LLET", respectively.

Stimuli and experimental procedures: As claimed by [5], [i:] is a long vowel when pronounced as HLET, MLET, and LLET. Hence, we selected three pairs of monosyllables with the rhymes [i:] and [i:t] from the entering and non-entering tone pairs. The three pairs of original sounds were taken from [6] to prepare the stimuli: (1) 枝[tsi:55] (branch) and 漪[tsi:t5] (squirt), with durations of 430 ms and 280 ms, respectively; (2) 智[tsi:33] (wise) and 节[tsi:t3] (festival), with durations of 450 ms and 300 ms, respectively; and (3) 字[tsi:22] (character) and 截[tsi:t2] (cut), with durations of 430 ms and 320 ms, respectively.

A total of 66 stimuli were prepared by manipulating the duration of each syllable in 10 steps with the Time-Domain Pitch-Synchronous Overlap-and-Add function in Praat [7]. The duration differences between stimuli in the same pairs were the same. In each experiment, the duration differences between stimuli were calculated from the duration differences between the two ends of the pair. For example, the durations of 枝[tsi:55] and 漪[tsi:t5] are 430 ms and 280 ms, respectively. A total of 20 stimuli were created, with 10 stimuli produced from 枝[tsi:55] and 10 stimuli produced from 漪[tsi:t5], each with a duration difference of (430 - 280) / 10 = 15 ms.

The experiments were conducted through Gorilla [8], an online experiment platform. The order of the three experiments was randomly determined by Gorilla, and each experiment took approximately five minutes to complete, with a five-minute break between experiments. During each experiment, the Gorilla system would randomly play 110 stimuli (22 stimuli, each repeated 5 times), totalling 330 trials of the 66 stimuli from the three experiments.

In each trial, the subjects first saw a "+" in the centre of the screen, which would disappear quickly. Then, a stimulus sound was played automatically, and two characters were presented on the page. The character on the left side was with a non-entering tone, and the one on the right was with an entering tone. After hearing the sound, the subjects were forced to choose between the two options. If the subjects thought they heard the character with a non-entering tone, they needed to press the "F" key with their left hand; otherwise, they should press the "J" key with their right hand.

Data processing and analysis: For each trial, "F" key selection was counted as 20% and "J" key selection as 0%. Since each stimulus was repeated 5 times, the result ranged from 0% to 100%. The data were analysed with a two-way ANOVA in R [9].

Results and discussion: In the three experiments, it was found that stop coda had a significant effect on the participants' judgments: $F(1,242) = 5434.92$, $p < .001$; $F(1,242) = 5173.42$, $p < .001$; $F(1,242) = 13076.38$, $p < .001$, while duration had no significant effect: $F(10,242) = 1.72$, $p > .05$; $F(10,242) = 1.18$, $p > .05$; $F(10,242) = 1.10$, $p > .05$. As illustrated in Figure 1, most of the results for the non-entering tones continuum were close to 100%, and most of the results for the entering tones continuum were close to 0%, suggesting that the perception of Guangzhou Cantonese monosyllabic entering tones was made according to the stop coda, regardless of the tone height. Only a small percentage of participants' judgments were influenced by duration. For instance, the

first nine stimuli for 智[tsi:33] were above 90%, and those of the tenth and eleventh stimuli were below 90%, indicating that when the duration of 智[tsi:33] is shortened, participants are more likely to perceive it as 节[tsi:t3]. Therefore, there are two findings on the perception of monosyllables with entering tone in Guangzhou Cantonese: (1) the stop coda is the distinctive feature; (2) the short duration is not a necessary feature, but it will influence perception. The findings suggested that the perception of entering tone in Guangzhou Cantonese is similar to that of Shanghai dialect [10] and Wu dialect [11], in which the stop coda is the distinctive feature. In addition, our findings provide experimental evidence supporting [2]'s suggestion that the stop coda is the distinctive feature of the entering tone perception and that the short duration is not necessary. Thus, the findings of this study contribute to a deeper understanding of the entering tones and phonological system of Cantonese.

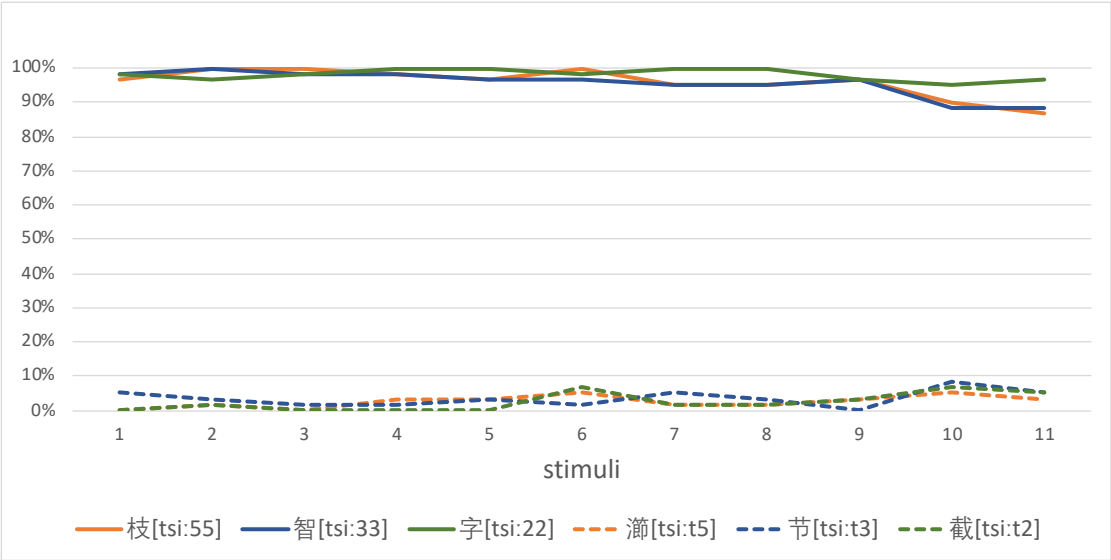

Fig.1 Means of 12 participants' judgements on stimuli.

References

[1] Liu, Y., Shi, F., Rong, R., & Sun, X. (2011). Xianggang yueyu shengdiao de fenzu fenxi [The tonal analysis of Hong Kong Cantonese in age groups]. *Yuyan Yanjiu,* (04), 98-106.

[2] Xia, Z. Y. (2006). Lun Rusheng duancu jishoucang - Rusheng lun zhi jiu [The 9th discussion of the Rusheng tone: Discussion of the feature of the Rusheng tone]. *Chengdu daxue xuebao - Sheke Ban,* (03), 88-90.

[3] Zhao, H. (1997). Qiantan hanyu Rushengyun seyinwei xiaoshi de yuanyin [A brief discussion on the reasons for the disappearance of final stop consonants in entering tone syllables in Chinese.]. *Guizhou Minzu Xueyuan Xuebao - Sheke Ban,* (02), 62-65.

[4] Zhu, X. N., Jiao, L., Yan, Z. C., & Hong, Y. (2008). Rusheng yanhua santu [Three ways of Rusheng sound change]. *Zhongguo Yuwen.* (4), 324-338.

[5] Shen, R. Q. (2015). Cong duanchang dao gaodi: Guangfupian yueyu Rusheng de shengxue xingzhi ji yanhua lujing [From short-long to high-low: The evolution of Cantonese stopped tones]. *In proceedings of the 18th International Conference on Yue Dialects.*

[6] Liu, X. Z. (2014). *Guangzhouhua danyinjie yutuce [Visible Cantonese: The spectrographic album of monosyllables of Guangzhou Cantonese].* World Publishing Corporation.

[7] Boersma, P. & Weenink, D. J. M. (2001). PRAAT, a system for doing phonetics by computer. *Glot International, 5*, 341-345.

[8] Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388-407. https://doi.org/10.3758/s13428-019-01237-x

[9] R Core Team. (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* URL https://www.R-project.org/.

[10] Gao, Y. F. (2004). *A study on tone perception* [Doctoral dissertation, Shanghai Normal University]. China National Knowledge Infrastructure.

[11] Yuan, D. (2015). *A study on phonetic variation in Wu dialect based on experimental analysis.* Shanghai People's Press.