

Vocabulary size predicts Mandarin-speaking older adults' perceptual learning of non-native tones

Kangdi Liu¹, Yan Feng², Quentin Zhen Qin¹

¹ The Hong Kong University of Science and Technology, ² Nanjing University of Science and Technology

While listeners can effortlessly perceive pitch variations in their native tones, non-native learners find it challenging [1]. Pitch contours serve as the primary cue for Mandarin tones, whereas Cantonese tonal contrasts rely on both pitch contours and pitch heights (e.g., T5-T6 is cued by pitch contours, hereafter, contour tones; T3-T6 is cued by pitch heights, hereafter, level tones) [2] (see Figure 1). Although Mandarin-speaking younger adults performed better in perceiving Cantonese contour tones than level tones overall during perceptual training, the degree of improvement was similar for both tones [3]. However, tone perception/learning declines with age [4]. Older adults face difficulties in processing dynamic pitch information [5], suggesting they might benefit from longer training. In addition, a large amount of individual variations was observed in older adults' speech learning due to age-related hearing declines [6]. Notably, older adults with larger vocabulary sizes tend to perceive more accurately at the initial stage through top-down mechanisms [7]. Based on the Scaffolding Theory of Aging and Cognition (STAC) [8], longer training may yield greater benefits for those with smaller vocabulary sizes. This study examines: how training improves older adults' accuracy in identifying Cantonese contour vs. level tones, given their age-related declines in tone perception (**RQ1**); whether vocabulary size predicts training gains in accuracy and whether this relationship varies by tone (**RQ2**).

Eighty-nine Mandarin-speaking older adults aged 55-78 were recruited, with relatively normal hearing and cognitive abilities. The auditory stimuli were novel spoken words produced by a native male speaker (see Figure 1). [ji] and [fən] carried contour tones, whereas [jəu] and [fu] carried level tones. Over five training phases with feedback, participants learned to associate these auditory stimuli with visual objects (see Figure 2). Each phase contained both words for every tonal contrast, repeated five times in total. Vocabulary sizes were assessed using accuracy on the Mandarin version of the Peabody Picture Vocabulary Test (PPVT) [9].

The mixed-effects model included Phase, Tone, and Vocabulary Size as fixed effects, with by-participant and by-item random effects. The analysis was conducted at the group level for RQ1 and at the individual level for RQ2. The results revealed a significant two-way interaction between Phase and Tone at the group level ($\beta = -0.17$, $SE = 0.02$, $z = -9.39$, $p < .001$), indicating that training improvement in older adults was more pronounced for contour tones (Contour: $\beta = -1.31$, $SE = 0.09$, $z = -14.66$, $p < .001$) than for level tones (Level: $\beta = -0.62$, $SE = 0.08$, $z = -8.18$, $p < .001$; see Figure 3A). A three-way interaction at the individual-level was found among Phase, Tone, and Vocabulary Size ($\beta = 0.11$, $SE = 0.02$, $z = 6.29$, $p < .001$). Post-hoc analyses showed that for contour tones only, participants with a lower vocabulary size gained more from the training ($\beta = 0.33$, $SE = 0.08$, $z = 4.15$, $p < .001$; see Figure 3B). This pattern was not observed for level tones ($\beta = -0.11$, $SE = 0.07$, $z = -1.69$, $p = .091$). It suggests that the role of vocabulary size in perceptual learning was modulated by the tone type.

The group-level results indicate that, in older adults, training led to greater improvement for contour tones than for level tones, unlike in younger adults [3]. Furthermore, our individual-level findings support the STAC [8]. Older adults with larger vocabularies employed top-down strategies in the word-object association task from the initial phase, resulting in higher baseline accuracy [7]. However, those with smaller vocabularies showed steeper learning slopes across training phases. The training was effective in compensating for their initially lower performance on Cantonese contour tones. It may indicate that Mandarin vocabulary knowledge scaffolds the learning of Cantonese contour tones. This study provides a deeper understanding of phonetic plasticity over the lifespan. It also creates an empirical foundation for developing personalized tone training paradigms to optimize learning outcomes.

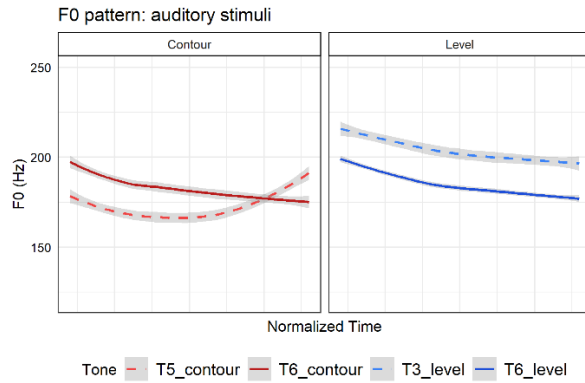


Fig. 1. The F0 pattern of Cantonese tones in stimuli. T5_contour and T6_contour (in red) represent the tonal contrast cued by pitch contours; T3_level and T6_level (in blue) represent the tonal contrast cued by pitch heights

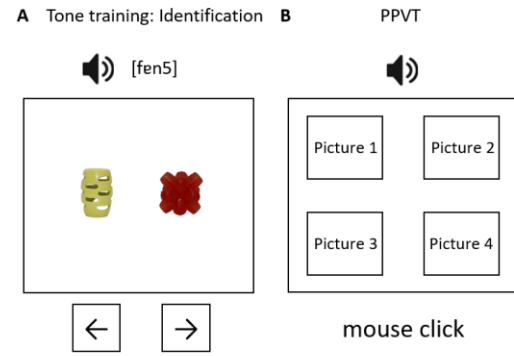


Fig. 2. Procedure of tone training (A, 80 trials in each phase: 2 tones $\times$ 5 tokens $\times$ 4 repetitions $\times$ 2 words; Counterbalance: /j/-initial and /f/-initial words; contour and level tones) and PPVT (B)

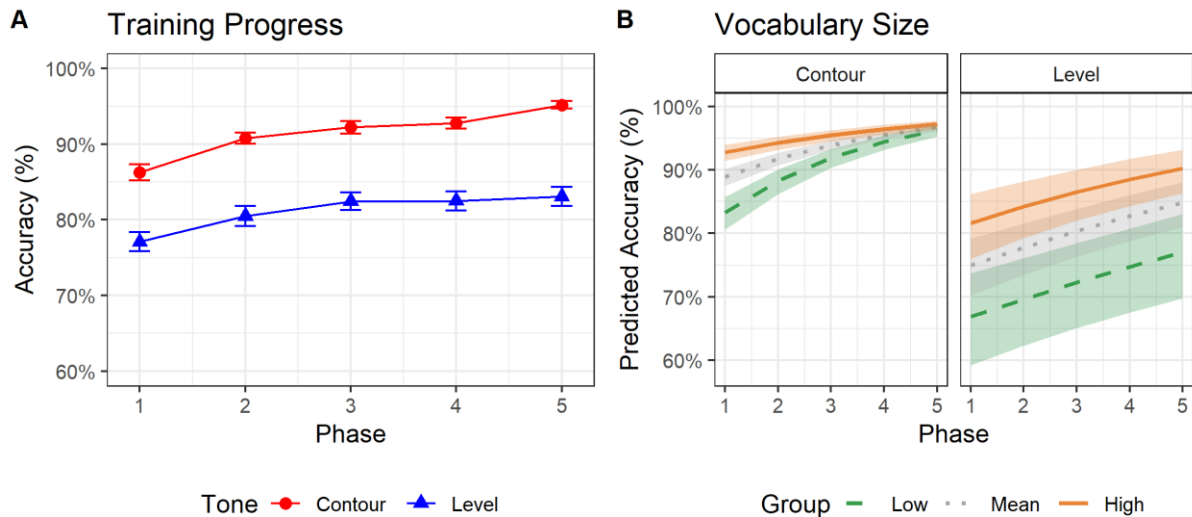


Fig. 3. Plot showing training progress by tone (A), and model-predicted plots showing the effects of vocabulary size (B) on training effects by tone (Note: Using the *ggpredict()* function, vocabulary size measurements were categorized into three groups: Low/Mean/High)

References

- [1] 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.
- [2] Qin, Z., & Jongman, A. (2015). Does second language experience modulate perception of tones in a third language? *Language and Speech*, 59(3), 318-338.
- [3] Qin, Q. Z., Jin, R., & Wu, R. (2025). The role of prior knowledge in second-language learners' overnight consolidation of Cantonese tones. *Journal of Phonetics*, 111, 101417.
- [4] Ingalson, E. M., Nowicki, C., Zong, A., & Wong, P. C. M. (2017). Non-native speech learning in older adults. *Frontiers in Psychology*, 8, 234801.
- [5] Shen, J., Wright, R., & Souza, P. E. (2016). On older listeners' ability to perceive dynamic pitch. *Journal of Speech, Language and Hearing Research*, 59(3), 1-11.
- [6] Shechter Shvartzman, L., Lavie, L., & Banai, K. (2022). Speech perception in older adults: An interplay of hearing, cognition, and learning? *Frontiers in Psychology*, 13, 816864.
- [7] Colby, S., Clayards, M., & Baum, S. (2018). The role of lexical status and individual differences for perceptual learning in younger and older adults. *Journal of Speech, Language, and Hearing Research*, 61(8), 1855-1874.
- [8] Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annual Review of Psychology*, 60, 173-196.
- [9] Tsoi, E. Y. L., Yang, W., Chan, A., & Kidd, E. (2019). Mandarin-English speaking bilingual and Mandarin speaking monolingual children's comprehension of relative clauses. *Applied Psycholinguistics*, 40(4), 933-964.