TALKER NORMALISATION OF PROSODIC CUES IN NON-NATIVE SPEAKERS

Bei Li (The Hong Kong Polytechnic University), Bruce Xiao Wang (The Hong Kong Polytechnic University), Si Chen (The Hong Kong Polytechnic University), Caicai Zhang (The Hong Kong Polytechnic University), Puiyin Lau (The University of Hong Kong) & Yike Yang (Hong Kong Shue Yan University)
benita.bei-li@connect.polyu.hk

In speech perception, the surface acoustic cues are susceptible to variability caused by social factors, as well as biological factors. Consequently, listeners must normalise the variable acoustic cues in speech perception. The process of listeners mapping variable acoustic signals onto the invariable underlying linguistic categories is called talker normalisation. Several talker normalisation models have been proposed (Chen et al., 2022; Shao & Zhang, 2019) and the current study aims to test the generalization of one of the latest talker normalisation models involving probabilistic parametric representation (PPR, Chen et al., 2022). Their model proposed that surface prosodic cues are represented as statistical distributions where the parameters of the distributions of the lexical tones are mentally stored, and this has been shown for both native Cantonese and Mandarin speakers in Cantonese lexical tone recognition (Chen et al., 2019, 2022). The mentally stored distributions may still play a role even if a tonal context is present for native Cantonese speakers. In the current study, we take tonal contexts into consideration in investigating the effect of tonal contexts and PPR on Cantonese tone recognition among Mandarin speakers.

Thirty-four Hong Kong Cantonese speakers (17 females and 17 males) were recruited and instructed to produce the target syllable /ji/ with T3 [33] embedded in two carrier sentences, i.e., /tʰiŋ tʰiŋ/ T1 [55] ("Listen to") __ and /tsɐu hɐi/ T6 [22] ("This is just") __. The tones of the carrier sentences were then manipulated based on the semitone distance (Table 1) between T1, T3 and T6 estimated from 68 speakers (CUSENT; Lee, n.d.). T1 [55] in /tʰiŋ tʰiŋ/ was manipulated to be 3.56 and 3.90 semitones higher than the target syllable (T3 [33]) for male and female speakers respectively; similarly, T6 [22] in /tsɐu hɐi/ was manipulated to be 3.56 and 3.90 semitones lower. Seven male and seven female native Mandarin speakers, with less than one-year formal Cantonese instruction, were recruited in the Cantonese tone recognition task where they were informed about the speaker gender in the stimuli. If the mentally stored distribution PPs play a more significant role in tone recognition, the target T3 would be identified as T3 regardless of its tonal context. On the contrary, if tonal context is more important, the target T3 would be identified as the low T6 in manipulated high T1 context and, *mutatis mutandis*, T3 would be identified as high T1 in manipulated low T6 context.

Multinomial mixed effects modelling was used to test the significance of tonal contexts and mentally stored distributions on Cantonese tone recognition among Mandarin speakers. The result shows that mentally stored distributions have a significant effect on Cantonese tone recognition in both manipulated low and high tone contexts ($ps < 0.001$), while only the low tone context has a significant effect ($p < 0.001$). Therefore, the findings provide insights about the accessibility of distributional information of non-native prosodic cues with the presence of contextual information.

	Semitone distance	T1 vs. T3	T1 vs. T6	T3 vs. T6
Table 1. Semitone distance between T1 and T3, T1 and T6 and T3 and T6 of male and female speakers from the CUSENT corpus.	**Male**	2.64	3.56	0.93
	Female	3.01	3.90	0.88

References

Chen, S., Zhang, C., Lau, P., Yang, Y., & Li, B. (2022). Modelling representations in speech normalization of prosodic cues. *Scientific Reports*, *12*(1), 14635. https://doi.org/10.1038/s41598-022-18838-w

Chen, S., Zhang, C., Lau, P., Yang, Y., Li, B., & Fung, W. K. (2019). A quantitative model of talker normalization by native and non-native speakers. *The Journal of the Acoustical Society of America*, *146*(4_Supplement), 2840. https://doi.org/10.1121/1.5136843

Lee, T. (n.d.). *Cantonese Spoken Language Data Resources* (1.0 - Continuous Cantonese sentences) [Data set]. http:// dsp.ee.cuhk.edu.hk/html/ cucorpora.html

Shao, J., & Zhang, C. (2019). Talker normalization in typical Cantonese-speaking listeners and congenital amusics: Evidence from event-related potentials. *NeuroImage. Clinical*, *23*, 101814.