Prosodic boundaries and givenness in tonal coarticulation in conversational Cantonese

Xin Gao[1], Cesko Voeten[1,2,3], and Mark Liberman[1]

[1]University of Pennsylvania (USA), [2]University of Amsterdam (Netherlands), [3]Fryske Akademy (Netherlands)

Introduction: This research investigates "Pre-low Raising" (PLR), where the pitch of the first syllable in a "high-low" sequence is realized higher than in a "high-high" sequence (Gu & Lee, 2009). While PLR is well-documented across languages in read speech (e.g., Lee et al., 2021), its presence in spontaneous speech is less understood. Our study examines PLR in conversational Hong Kong Cantonese, focusing on the effects of prosodic factors and information structure, specifically boundaries and givenness.

Existing research suggests coarticulation is influenced by prosodic boundaries (Li & Chen, 2019) and givenness (Baker & Bradlow, 2007). Given that prosodic boundaries can weaken or even block coarticulation, we hypothesize in conversational Cantonese, prosodic boundaries will reduce the PLR effect. Additionally, we predict PLR realization will differ between first (i.e. new) and subsequent (given) mentions.

Methods: This study examined the pitch realization of Tone 1 (T1), a high-level tone, in contexts where it precedes either another T1 or a low-register Tone 4 (T4). This involves comparing the F0 in the first syllable of the T1-T1 and T1-T4 sequences.[1] We used CantoMap, a corpus featuring approximately 13 hours of Hong Kong Cantonese conversational data from 40 speakers, complete with prosodic boundary information (Winterstein et al., 2020). F0 measurements from the T1-T1 and T1-T4 sequences were taken at every 10% interval of sonorant segments and normalized for each speaker.

We fit a generalized additive mixed model with fixed effects and smooth terms for time by second-syllable tone, boundary (present/absent), and givenness (first mention/later mention). We also included smooths for speech rate and the word location within sentence. By-speaker random effects were added for all terms. Based on exploratory fits, the model was fit to scaled-*t* errors and included an autoregressive process of order 1 with ρ = .3. Significance of differences between conditions was assessed using Bayesian credible intervals. Differences are considered significant wherever the credible interval excludes zero.

Results: **Fig. 1** presents F0 contours for T1-T1 and T1-T4 sequences, distinguishing first and second syllables' pitch (left and right of the dashed line, respectively). It suggests that pitch is higher in T1-T4's first syllable over T1-T1, evidencing PLR. **Figs. 2-4** present the relevant difference curves computed from the GAMM fit. **Fig. 2** illustrates the F0 difference in the first syllable between T1-T1 and T1-T4, with negative y-axis values indicating lower F0 in T1-T1.

Boundary presence significantly affects PLR: no inter-syllable boundary increases F0 difference in the first syllable of T1-T1 and T1-T4 sequences (**Fig. 3**). However, givenness of the phrase doesn't show significant influence on PLR, with the confidence interval for differences always including 0 (**Fig. 4)**.

Discussion: PLR, a dissimilatory coarticulation affecting phonetic realization based on the phonological tonal categories, is not merely due to articulatory gesture overlapping. These results demonstrate the presence of PLR in spontaneous speech.

The current study highlights the roles of prosodic structure in speech coarticulation. Specifically, the findings indicated that PLR is stronger in the absence of a prosodic boundary, indicating the influence of prosodic boundaries on tonal coarticulation.

This result provides evidence that the speech production is a complex interplay of articulation, phonology, and prosodic structure, and emphasizes the need to incorporate all these elements in speech articulation and coarticulation studies. Furthermore, spontaneous speech in real-world contexts is distinct from the simplified nature of read or scripted speeches, which lack the complexity of message structures.

---

[1] T6 is also a low-register tone in Cantonese; T1-T6 sequences were not analyzed due to the frequent occurrence of the word *gan1 zyu6* "and then" in the corpora, primarily in utterance-initial positions.

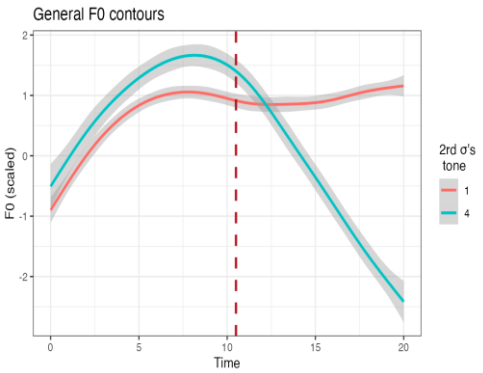

Fig. 1: F0 contours, distinguished by second syllable's tone category.

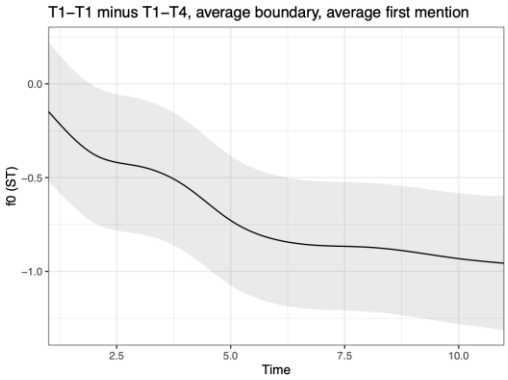

Fig. 2: Difference in first syllable's F0 between T1-T1 and T1-T4 sequences

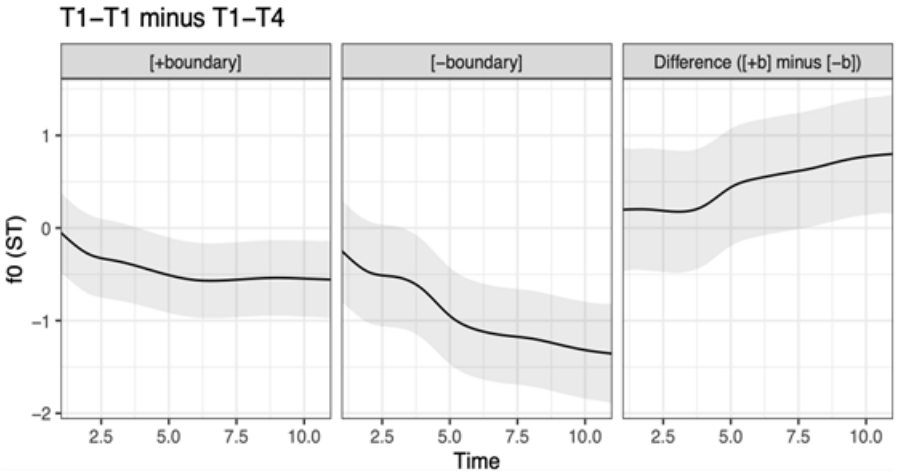

Fig. 3: Boundary effect on F0 difference between T1-T1 and T1-T4 sequences

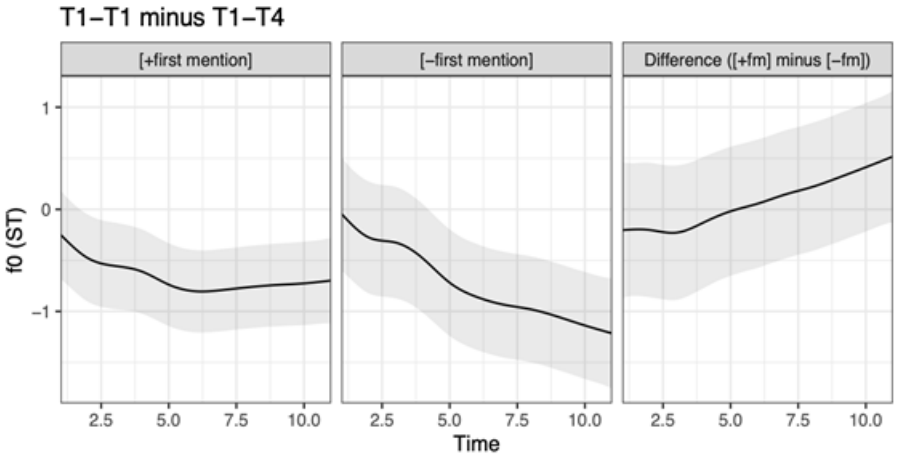

Fig. 4: Givenness effect on F0 difference between T1-T1 and T1-T4 sequences

References

[1] Baker, R. E., & Bradlow, A. R. (2007). Second mention reduction in Indian, English, and Korean. Journal of the Acoustical Society of America, 122(5), 2993.

[2] Gu, W., & Lee, T. (2009). Effects of tone and emphatic focus on F0 contours of Cantonese speech: A comparison with Standard Chinese. Chinese Journal of Phonetics, 2, 133-147.

[3] Lee, A., Prom-On, S., & Xu, Y. (2021). Pre-low raising in Cantonese and Thai: Effects of speech rate and vowel quantity. The Journal of the Acoustical Society of America, 149(1), 179-190.

[4] Li, Q., & Chen, Y. (2019). Prosodically conditioned neutral-tone realization in Tianjin Mandarin. Journal of East Asian Linguistics, 28, 211-242.

[5] Winterstein, G., Tang, C., & Lai, R. (2020, May). CantoMap: a Hong Kong Cantonese MapTask Corpus. In Proceedings of the Twelfth Language Resources and Evaluation Conference (pp. 2906-2913).