Individual variation in the prosody of Cantonese rhetorical questions

Roger Yu-Hsiang Lo[1], Angelika Kiss[2], and Maxime Tulling[3]
[1]University of British Columbia, [2]University of Toronto, [3]New York University
[1]roger.lo@ubc.ca, [2]angelika.kiss@mail.utoronto.ca, [3]mat646@nyu.edu

Background: Due to the constraints imposed by the rich lexical tone inventory of Cantonese, it has been argued that the locus of intonation is on the final syllable, often occupied by a sentence-final particle (SFP) (e.g., Xu and Mok, 2011). Indeed, Lo et al. (2019) showed that, in production, the fundamental frequency (F0) and duration of SFPs alone are sufficient to distinguish different readings of a string-identical question. However, other studies have reported a global F0 raising effect that can be observed in question intonation (Tsui and Tong, 2018). We are interested in whether prosodic marking between different question types in Cantonese is also already present before the right boundary of the utterance, and whether individuals differ in this aspect. We considered the F0 contour and duration of *wh*-questions, such as *Who likes liquorice?*, produced in contexts supporting an information-seeking (ISQ), a negative rhetorical (RQ−, where the suggested answer is 'nobody'), or a positive rhetorical reading (RQ+, where the suggested answer is a certain person).

Experiment: We created 12 target *wh*-questions (all with the *wh*-word *bin1go3* 'who') that ended with either the SFP *aa1* 吖 or *aa3* 呀. We included both, since there is no standardized convention for writing SFPs in Cantonese, and these two SFPs seem to be used interchangeably. Crucially, these questions share the same syntactic structure and the same number of syllables, and are ambiguous among ISQ, RQ− and RQ+ readings. Each target question was embedded in three short contexts, with each context favoring one of the three readings. Twenty-two native speakers of Cantonese participated in the experiment, and were tasked to read aloud the target questions naturally after listening to the contexts. We excluded the data from seven speakers due to difficulty in producing casual Cantonese from written forms and inattentiveness during the experiment.

Analysis: We opted for Functional Principal Component Analysis (FPCA, Gubian et al., 2015) to analyze F0 contours, as FPCA does not require the researcher to define a priori the regions of interest for statistical modeling and offers a way to succinctly quantify global variation in F0. In essence, FPCA treats each F0 contour as a composition of a grand mean curve and a small number of Principal Component (PC) curves (each of which encodes a different deformation of the mean curve) weighted by corresponding PC scores (which can be studied using conventional statistical tests). We extracted the F0 contour and the duration of each syllable in the utterance using Praat (Boersma and Weenink, 2019). F0 contours were then subjected to the FPCA, and the first two PC scores were used for further analyses. Because the results from questions ending in *aa1* follow the same pattern as those from *aa3*, we conflated these two SFPs in our analysis.

Results: The effect of the first two PC curves (i.e., PC1 and PC2) on the mean F0 curve is displayed in Figure 1. Focusing on the *wh*-word and the SFP (because they are located at the beginning and end of the utterance), PC1 coordinates both the F0 height on the *wh*-word and the SFP, while PC2 mainly alters the pitch level of the SFP. The reconstructed F0 contours based on these two PCs and the relative duration of the *wh*-word and the SFP to the utterance duration for each speaker are shown in Figure 2a and 2b. To investigate the effect of the question type on F0 and duration, we then ran a class analysis to examine if question types are correlated with PC scores and (relative) duration of the *wh*-word and the SFP (e.g., $\text{PC1 Score} \sim \beta_0 + \beta_1 \cdot \text{Question Type}$). The results indicate that, at the group level, the question type is only predictive of PC1 scores and duration of the SFP, but not PC2 scores or duration of the *wh*-word. At the individual level, however,

speakers do not always conform to this observation. In sum, our findings suggest that, overall, prosodic marking of question types occurs prior to the last syllable in terms of F0, but not in terms of duration. However, this pattern does not apply to each speaker; differences might be attributed to variation in individual grammars or an artifact of non-spontaneous production in a lab setting.

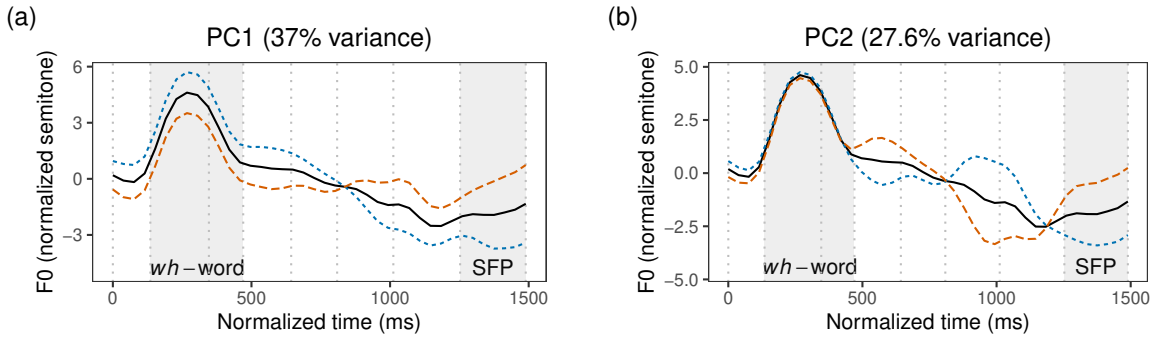

Figure 1: Each panel shows in solid the mean curve $\mu(t)$ and the dashed lines obtained by adding (orange) or subtracting (blue) from $\mu(t)$ the curve (a) $\sigma(\text{PC1 scores}) \cdot \text{PC1}(t)$ and (b) $\sigma(\text{PC2 scores}) \cdot \text{PC2}(t)$ respectively, where σ denotes the standard deviation.

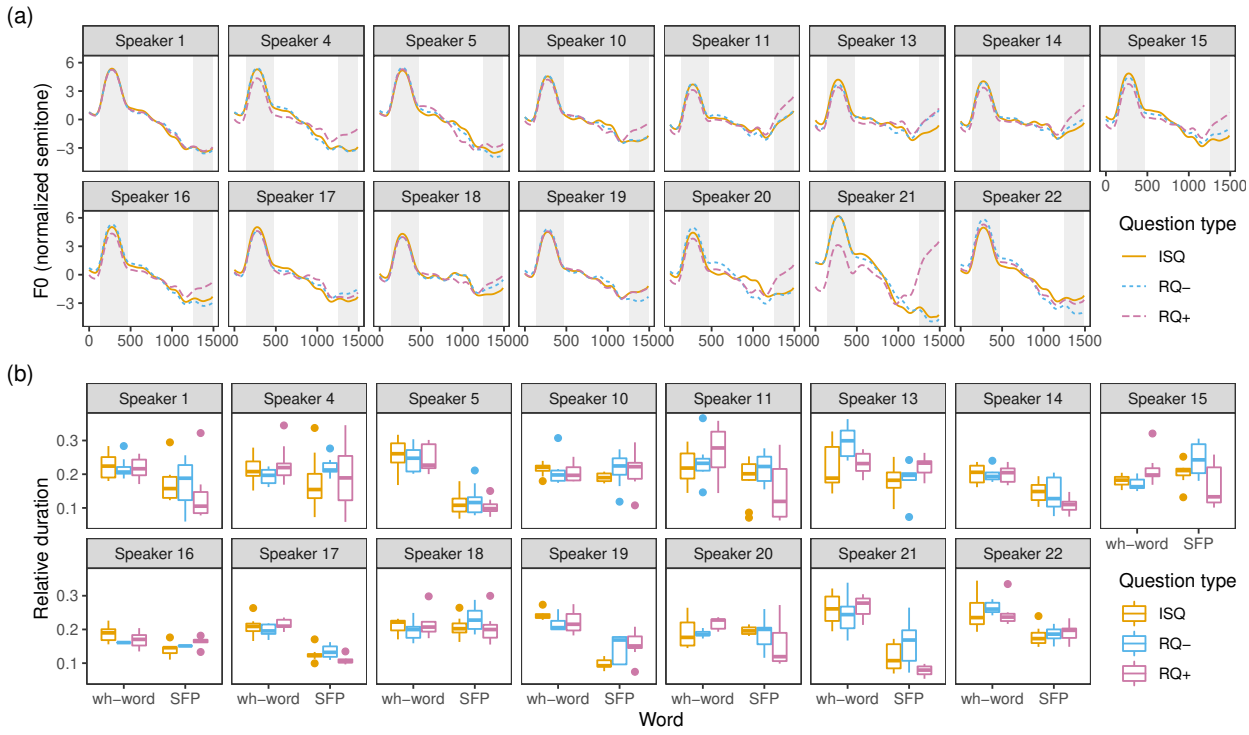

Figure 2: Individual variation in (a) reconstructed F0 contours from PC1 and PC2 curves and in (b) relative duration of the *wh*-word and the SFP to the utterance duration.

References

Boersma, P. and Weenink, D. (2019). Praat: Doing phonetics by computer [computer program].

Gubian, M., Torreira, F., and Boves, L. (2015). Using Functional Data Analysis for investigating multidimensional dynamic phonetic contrasts. *J. Phon.*, 49:16–40.

Lo, R. Y.-H., Kiss, A., and Tulling, M. A. (2019). The prosodic properties of the Cantonese sentence-final particles *aa1* and *aa3* in rhetorical *wh*-questions. In *Proc. ICPhS 19*, pages 502–506, Melbourne.

Tsui, R. and Tong, S. (2018). An acoustic study of the effect of the interaction between intonation and lexical tones on the realization of Cantonese sentence-final particles. In *Proc. Speech Prosody 9*, pages 45–49, Poznań.

Xu, B. R. and Mok, P. (2011). Final rising and global raising in Cantonese intonation. In *Proc. ICPhS 17*, pages 2173–2176, Hong Kong.