Acoustic correlates of question-final Cantonese tones

Una Y. Chow
The University of British Columbia

Fundamental frequency (f0) has been long reported as the primary acoustic correlate of lexical tones in Cantonese (e.g., Vance, 1976). The six lexical tones in Hong Kong Cantonese (Bauer & Benedict, 1997) can be characterized by f0 height (high-level, mid-level, and low-level) and/or f0 direction (high-rising, low-rising, and low-falling). Khouw & Ciocca (2007) performed a discriminant analysis on the f0 measurements of eight consecutive sections of the vocalic portion of 24 monosyllabic words, read by 10 Cantonese speakers. The analysis revealed that magnitude of f0 change over the 6th and 7th sections of the tone was important in identifying the tones.

Khouw & Ciocca (2007) examined the Cantonese tones in isolated words. Tones tend to resemble their canonical patterns when produced in isolated words but will likely differ from these patterns when produced in continuous speech due to factors such as co-articulation and intonational effects. For example, declarative questions in Cantonese (e.g., *Jyu4 So2 hai6 jat1 go3 ngaa4 ji1*? 'Jyu So is a dentist?') typically end with a high-rising intonation regardless of the shape of the final tone (Wong et al., 2005). Consequently, listeners are likely to misperceive the other tones in question-final position as the high-rising tone (e.g., mid-level, low-level, low-rising, and low-falling tones in Ma et al., 2006).

This study aimed to address the research questions: (1) To what extent can f0 height and magnitude of f0 change serve as acoustic cues to Cantonese tones in utterance-final syllables of declarative questions? (2) For each of these two cues, which portion of the tone is most important in identifying the Cantonese tones?

The declarative questions for the current analysis were collected for a study on Cantonese intonation. Ten native Hong Kong Cantonese speakers (5 male, 5 female; age 21–26 years) read twice 20 pairs of statements and declarative questions arranged in five blocks. In each block, the four final syllables consisted of the same segments (e.g., si) but each carried a different tone. In the acoustic analysis, the low-rising tones were combined with the high-rising tones and the mid-level tones were combined with the low-level tones because the paired tones are merging for some speakers (Mok et al., 2013); Whitehill et al. (2000) also found no significant difference in the overall f0 between the mid-level and low-level tones. Combining the tones results in four tonal groups: H (high-level), M/L (mid/low-level), R (low/high-rising), and F (low-falling).

Final syllables were obtained from the 20 declarative questions (10 speakers × 5 syllables × 4 tonal groups × 2 readings = 400 tokens). Then a Praat script (Boersma & Weenink, 2019) extracted the two acoustic measures: f0 height (or mean f0 in semitone) at 10 equal intervals of the periodicity of the rhyme of the final syllable (i1, i2, i3, ..., i10) and magnitude of f0 change (or the absolute value of the mean f0 difference) between consecutive intervals (i1-i2, i2-i3, i3-i4, ..., i9-i10). For each acoustic measure, a conditional random forest model (*cforest* in R; R Core Team, 2019) was trained and tested on the interval values in classifying the four tonal groups. A 2-fold cross-validation was performed, with tokens from the first and second readings, in turn, served as training and test data, respectively.

Figure 1 displays the f0 height and magnitude of f0 change by tonal group. In both cross-validation test runs, the *cforest* model ranked i4 for f0 height and i6-i7 for magnitude of f0 change as the most important part of the final tone in classifying the Cantonese tones. The latter is consistent with Khouw & Ciocca's (2007) result, demonstrating the robustness of the magnitude of f0 change cue at the later part of the tone. As expected, the important part of the tone for f0 height is earlier, due to the effect of the question rise on the later part of the tone. The mean accuracy scores (in Cohen's Kappa values) of the two test runs for f0 height (0.48665) and magnitude of f0 change (0.48) are nearly the same; however, magnitude of f0 change may be more robust because the important part of the tone for f0 change is earlier, which can be affected by the onset or preceding syllable. Sensitivity was also computed based on the rate of correctly identified tones. The model was most sensitive to the high-level tone; this tone is already high so it is less affected by the question rise. The model was also fairly sensitive to the low-falling tone, which dips lower than the other tones. However, the model was much less sensitive to the mid/low-level tones and least sensitive to the low/high-rising tones due to the effect of the question rise on these tones. Non-f0 cues would be necessary to distinguish them.

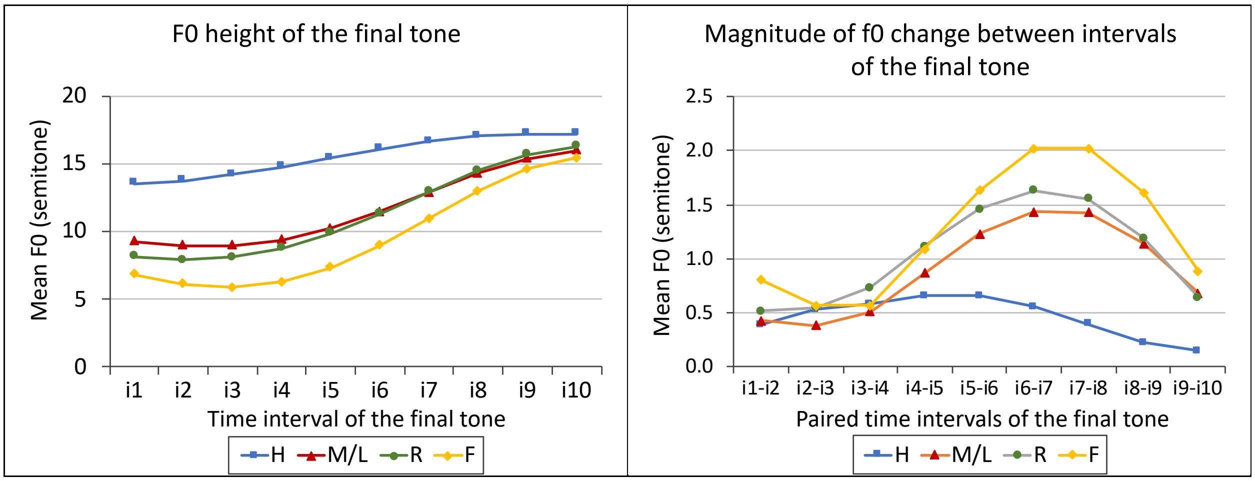

Figure 1: F0 height and magnitude of f0 change of the final tone, by tonal groups

References

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese Phonology*. Mouton de Gruyter.

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 6.1.05). Retrieved from http://www.praat.org

Khouw, E., & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics*, *35*, 104–117.

Ma, J. K., Ciocca, V., & Whitehill, T. L. (2006). Effect of intonation on Cantonese lexical tones. *The Journal of the Acoustical Society of America*, *120*(6), 3978–3987.

Mok, P. P., Zuo, D., & Wong, P. W. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change*, *25*(3), 341–370.

R Core Team. (2019). R: A language and environment for statistical computing (Version 3.6.1). R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Vance, T. J. (1976). An experimental investigation of tone and intonation in Cantonese. *Phonetica*, *33*, 368–392.

Whitehill, T. L., Ciocca, V., & Chow, T. Y. (2000). Acoustic analysis of lexical tone contrast in dysarthria. *Journal of Medical Speech—Language Pathology*, *8*, 337–344.

Wong, W. Y. P., Chan, M. K. M., & Beckman, M. E. (2005). An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In S. A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 271–300). Oxford University Press.