IVPAI 2026

The 8th International Conference on Image, Video Processing and Artificial Intelligence

Pitch Features of Disyllabic Dunyu in Cantonese Reading Discourse

Paper ID： VP811

Dechong Deng

School of Culture and Communication, Guangdong Polytechnic of Science and Technology, Zhuhai, Guangdong, China

E-mail: janeyddc@163.com

Abstract

This study examines the pitch features of disyllabic Dunyu (prosodic phrases) in Cantonese reading discourse. Based on a six-level prosodic hierarchy, F0 was extracted and converted to semitones across four genres. Results show that disyllabic Dunyu generally exhibit a declination trend, though ~18% display a final-syllable rise. Syllable position significantly affects pitch, with the first syllable consistently higher than the second. Crucially, this position effect is tone-dependent: minimal for the high-level tone (T1) but highly pronounced for the low-rising tone (T5). Additionally, pitch drop before the final syllable is modulated by pause boundaries, being significantly deeper before major (Level-1) pauses than minor (Level-2) ones. This study reveals the pitch organization rules of Cantonese disyllabic Dunyu, providing new empirical evidence for Cantonese intonation research and offering insights for speech synthesis and pedagogical practice.

Methods/Results

Methods

- Data: 8 native speakers; 4 genres of Cantonese reading discourse.
- Processing: F0 extracted via Praat, converted to Semitones (St) relative to 75 Hz.
- Segmentation: Douju bounded by major pauses (>200ms); Disyllabic Dunyu extracted within Douju via minor pauses (80–200ms).

Results

- Overall Trend: Steep declination in the first 60%, then flattens. ~18% show a final V-shape rise in interrogative/continuative contexts.
- Position Effect: Robust local declination; the 2nd syllable is consistently lower than the 1st.
- Tone Difference: T1 is highly stable and resistant to declination. T5 is anomalous, showing no final lowering and a wider 2nd-syllable range for tonal compensation.
- Boundary & Domain: Pitch range expands differently by Douju position. Major pauses induce deeper final lowering, but this effect is blocked in T5 to preserve its identity.

Figures

Figure 1: Overall Pitch Contour of Disyllabic Dunyu

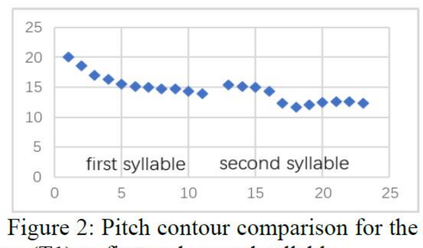

Figure 2: Pitch contour comparison for the high-level tone (T1) as first and second syllable.

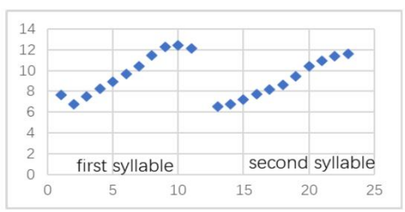

Figure 3: Pitch contour comparison for the high-rising tone (T2) as first and second syllable.

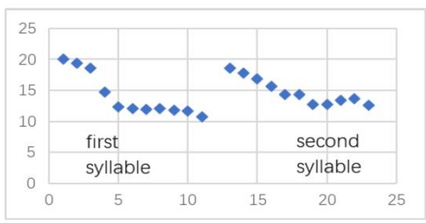

Figure 4: Pitch contour comparison for the mid-level tone (T3) as first and second syllable.

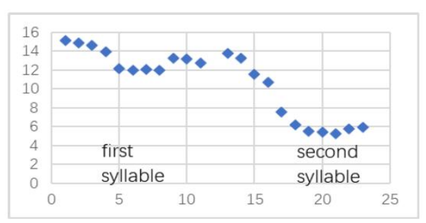

Figure 5: Pitch contour comparison for the low-falling tone (T4) as first and second syllable.

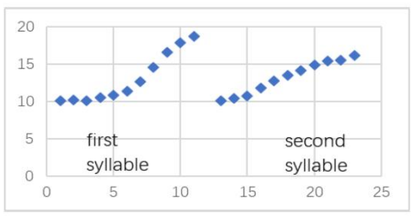

Figure 6: Pitch contour comparison for the low-rising tone (T5) as first and second syllable.

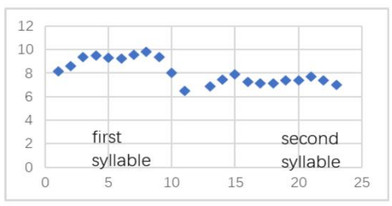

Figure 7: Pitch contour comparison for the low-level tone (T6) as first and second syllable.

Conclusions

In conclusion, this study establishes that the pitch contour of Cantonese disyllabic Dunyu is governed by a set of interrelated factors: a fundamental declination trend, a strong effect of syllable position that interacts with tone-specific properties, a Douju-positional effect on pitch range that distinguishes Cantonese from Mandarin, and a boundary-induced lowering that is modulated by tone category. The low-rising tone (T5) consistently emerges as a distinctive element within this system. These results offer robust empirical support for a hierarchical, tone-sensitive model of Cantonese prosody and provide a valuable foundation for future work in speech technology and language education. Subsequent research should extend this inquiry to spontaneous speech and diverse speaker populations to further test the generalizability of these patterns.