Cantonese rising tone alignment as a phonological regularity

Introduction. The study of how fundamental frequency contour (f0) aligns with segments has important implications for not only description of tonal and intonational systems, but also theories of intonational phonology. For example, Xu (1998) found that the f0 peak of the rising tone in Mandarin consistently occurs after the offset of the tone-carrying syllable regardless of speaking rate and whether the syllable has a final nasal. This has been taken as evidence that tone is fully syllable-synchronised, with the peak delay attributed to inertia. This tone-syllable synchrony becomes a foundational assumption of the Time Structure model of the syllable (Xu & Liu, 2006) and PENTA model of intonation (Xu, 2005). However, it remains unclear whether this synchrony can be extended to other tone languages or would autosegmental anchoring of tone targets instead be observed (e.g. Arvaniti et al. 1998). Here we report an experiment on the effect of segmental composition on f0 peak alignment of the two rising tones (i.e. T2 [25] and T5 [23]) in Cantonese.

Method. Meaningful monosyllables /ji:/, /ŋa:/, /ji:m/, and /si:/ were concatenated to form three types of disyllabic non-words—/ji: ŋa:/, /ji:m ji:/, and /ji: si:/. 9 native speakers of Hong Kong Cantonese (5m and 4f, aged 19-28) read aloud these disyllables in a carrier sentence *"我講xx俾你聽"* ("I say xx to you"). To control for speaking rate, regular beats were played at an interval of 3s and subjects were instructed to read aloud each sentence between two beats.

Results. Segmental composition affects f0 peak alignment of lexical tones in a way reminiscent of intonational tones (D'Imperio, 2000). Figures 1 to 3 show the f0 contour in three types of disyllabic non-words, in which the first syllable carries one of the two rising tones. Except when followed by the high level tone, the f0 peaks of the two rising tones predominantly occur on the first syllable for /ji:m ji:/, but either close to the syllable boundary or on the second syllable for /ji: ŋa:/. In addition to consistency across following tone types, peak placement is highly regular across and within speakers and appears lawfully governed.

Discussion. No strict segmental anchoring of f0 peak is observed, and its absence cannot be attributed to differential time pressure on the first syllable (Ladd et al, 1999; Schepman et al., 2006) as the two rising tones, despite their distinct slopes, have almost identical peak locations across syllable compositions. We therefore hypothesise that Cantonese rising tone alignment may be regulated at a higher level than segments, with the peak specified as a percentage of the feet. Quantitative evidence of the correlation between feet duration and f0 peak location is used to corroborate this hypothesis. Moreover, unlike Mandarin (Xu, 1998), the peak of the Cantonese rising tones may occur well before the syllable boundary (e.g. /jim: ji/); these early peaks cannot be the product of inertia and must be phonologically specified as such. This weighs against tone-syllable synchrony as a universal principle and therefore the PENTA model of intonational phonology. Finally, the lack of consistent alignment of f0 peak with reference to syllable boundary casts doubt on the usefulness of classifying tone languages based on the relative degree of carry-over and anticipatory coarticulation, e.g. the same tonal sequence in Cantonese can show more anticipatory coarticulation in /ji:m ji/ than in /ji: ŋa:/. Future studies on tonal coarticulation should systematically vary segmental compositions to fully explore potential phonological regularity at higher prosodic levels.

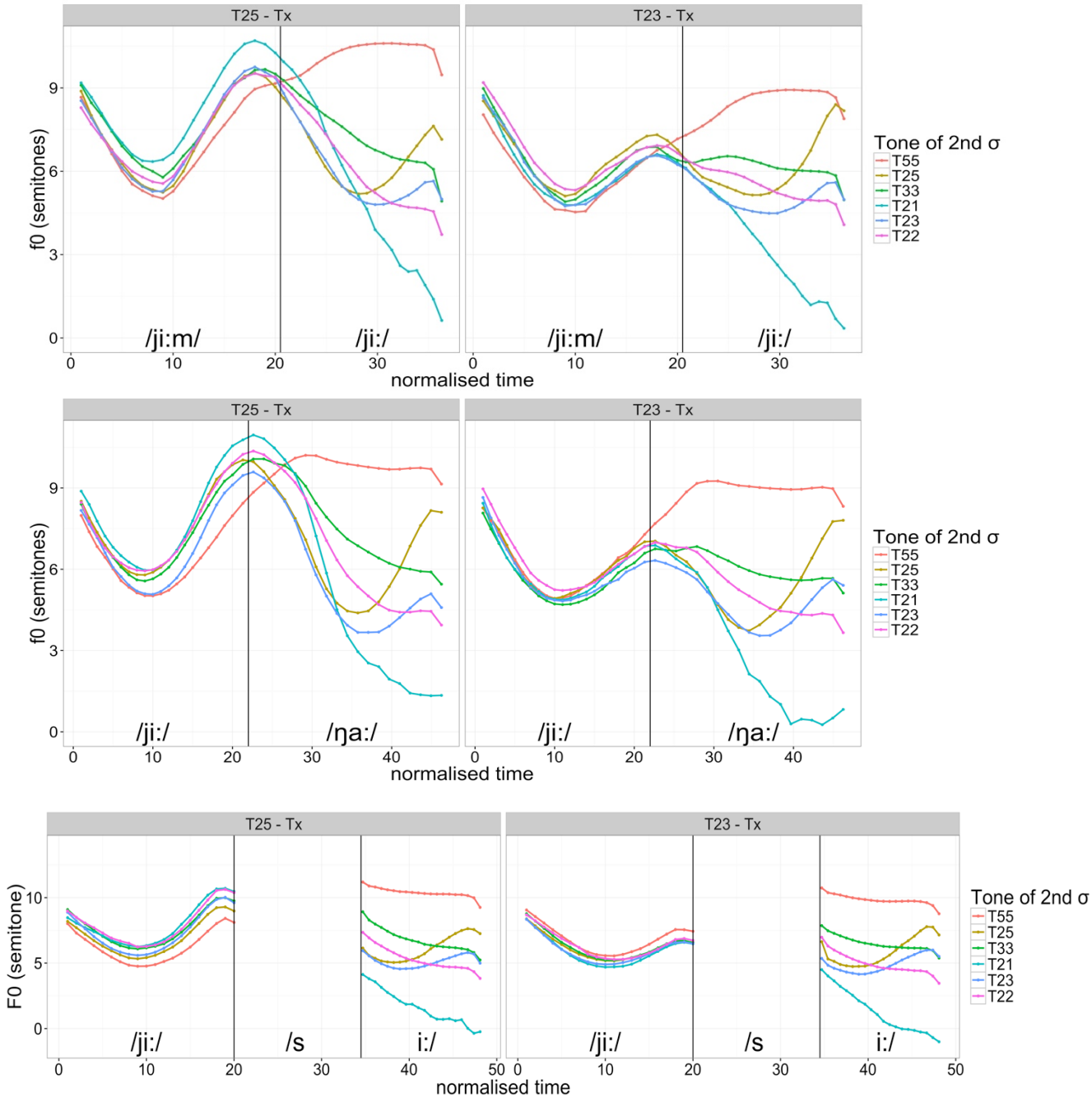

Figure 1 to 3. f0 contours of T2[25]-Tx (left panels) and T5[23]-Tx (right panels) disyllabic non-words /ji:m ji:/, /ji: ŋa:/, and /ji: si:/ respectively. X-axis depicts the normalised durational ratio of the segments; y-axis represents the raw f0 values in semitones. Vertical lines represent the location of the syllable boundary (in Figure 1 and 2) and segmental boundary (in Figure 3). The f0 contours of /s/ in Figure 3 are removed as the estimate is unreliable.