"Rising" Intonation on "Falling" Tones

Masayuki Gibson
Cornell University

The nature of the interaction between sentence level intonation and lexical tone varies from language to language. This is clearly evident in how "rising" intonation (an intonational contour that is perceived as rising) interacts with the lexical tones on words near the right edge of the utterance in different languages. Adherents of the ToBI-style notation (Pierrehumbert and Beckman 1988, e.g.) analyze such utterances as bearing a H boundary tone at the right edge. While this is a reasonable analysis, it is a necessarily language-dependent one that is insufficient to capture the typological variation that is apparent upon inspection of data from multiple languages. Yi Xu's (2005) PENTA model treats tones and intonation as separate functions that are implemented by the Phonetics in parallel. This model assumes that tone and intonation do not interact in the Phonology. Results from production and perception experiments that were conducted in several languages for this study, including Shiga Japanese, Mandarin, Cantonese, and North Kyeongsang Korean suggest that any model of speech melody must allow for both language-specific phonetic implementation and for the interaction of tone and intonation in the Phonology.

Shiga Japanese: Unlike in Tokyo Japanese, finally-accented words in SJ retain the drop in pitch associated with the accent on a final light syllable, thus maintaining the contrast between finally-accented and unaccented words. The realization of the pitch drop in this dialect does not seem to require a lengthening of that final mora/syllable. Meanwhile, when pronounced with an echo question intonation, the final mora of a finally-accented word still displays a sharp rise *after* the drop associated with the accent. The realization of this rise is accompanied by a drastic lengthening of the last mora (doubling its length in most cases). (See Figure 1a)

Mandarin: Echo questions in Mandarin (Putonghua) are characterized by a raising of the overall pitch level that causes a final lexical falling tone to be in a higher range than in a declarative utterance (but falling nonetheless). Yuan (2004) shows that in echo questions the F0 is shifted up from the start of the utterance and that the pitch range of the final syllable is increased. Results from the present production study confirm this. Yuan's observation that the effect of the echo question intonation on the tones is tone-specific is also corroborated; the first and second tones seem simply to be shifted upwards, whereas the third tone gets pulled down just as low as a declarative third tone but then ends much higher and the beginning of the fourth tone gets shifted upwards to a greater degree than the end. Overall, we don't see duration effects in Mandarin that are comparable to those seen in SJ, but we do see a slight lengthening of the final syllable for one speaker that could be attributed to phonetic marking of focus (see Chen 2002). (See Figure 1b)

Cantonese: In Hong Kong Cantonese, the intonational rise of an echo question is realized on the final syllable. The results of the present study show that, unlike in Shiga Japanese, a final falling tone in Cantonese does not complete its fall before the rise is initiated (c.f. Wu 1990, who reports that "the rise starts after the fall" and Yip 2002, who claims that the tone starts where it would in a declarative context and ends "high"). Like in Mandarin, we see some tone-specific effects. Also like in Mandarin, the duration of the final syllable is not affected to the degree that we see in SJ. (See Figure 1c)

North Kyeongsang Korean: Despite its status as a so-called "pitch accent" language, NKK behaves quite differently from Japanese when it comes to the reconciliation of a lexical HL sequence with a rising intonation. If the HL sequence falls on the last two syllables of an echo question, we still see a lower F0 on the second syllable, though it doesn't drop as low as it does in a declarative context. This is similar to the Mandarin case. However, if the HL sequence falls on a single syllable, there is no fall; the pitch simply rises and keeps rising. (See Figure 1d)

An adequate model of speech melody must minimally include a phonological component and a phonetic component. The phonological component must be able to do "repairs" like tone deletion (to handle NKK) and TBU lengthening (to handle SJ). The "rising" intonation scheme must also be sensitive to the tonal category (to handle the tone-specific phonetic implementation in Cantonese and Mandarin), rendering parallel implementation of tones and intonation impossible to maintain. Such a model would not only give us better descriptive power but would also go further than the other models mentioned above in accounting for cross-linguistic differences in perception. For example, Cantonese speakers are better able to identify the sentence type of an echo question but not as good at identifying the final lexical tone, whereas the reverse is true for Mandarin speakers. This asymmetry is likely due to the fact that, while the tone-specific implementation of intonation reinforces differences among the tones in Mandarin, it nearly neutralizes several of the tones in Cantonese.
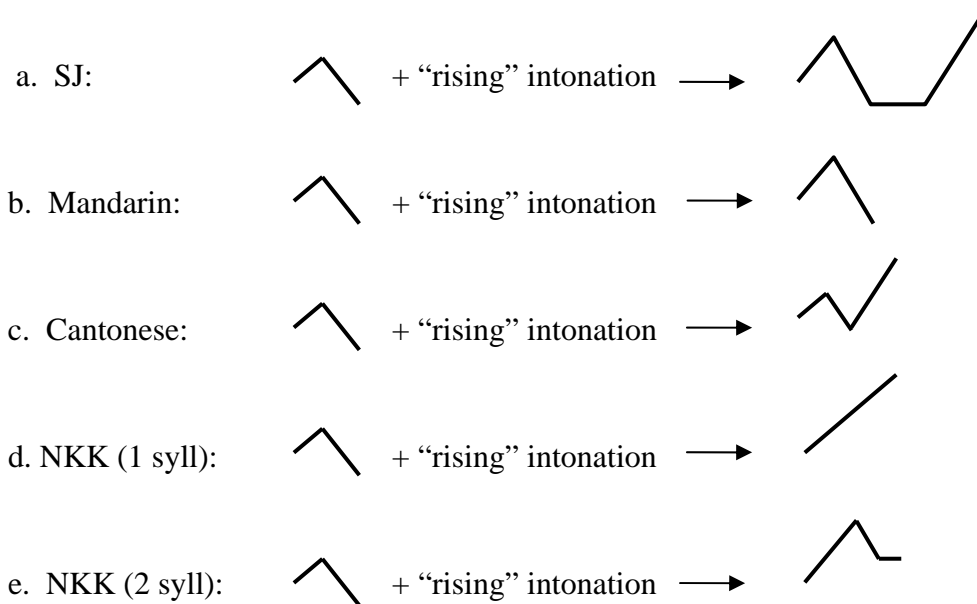

Figure 1: A schematic representation of the interaction of lexical HL tones with "rising" intonation associated with echo questions in (a) Shiga Japanese, (b) Mandarin, (c) Cantonese, (d,e) North Kyeonsang Korean.