Effects of anticipatory dissimilation on the F0 and alignment of Thai contour tones

In addition to three level tones, Thai contrasts two contour tones, which are restricted to bimoraic words. They are true contours, with pitch extrema ('elbows') near the middle of the word: the Falling tone (HL) rises before falling, while the Rising tone (LH) falls before rising. Morén and Zsiga (2006) link the contour shape and distribution to a moraic tone-bearing unit, arguing that tones associate to the right edge of the mora. This approach maintains that tones are driven by segments, and predicts that changes in tone timing parallel changes in segmental timing. I present results from an acoustic study on Thai tone that support an alternative hypothesis: lexical tones interact with each other and generate timing independently of segments.

Previous studies on Thai coarticulation (Gandour, Potisuk, Dechongkit, and Ponglorpisit, 1992; Potisuk, Gandour, and Harper, 1997) have argued that anticipatory effects are largely dissimilatory and limited to the latter portions of tones; however, they primarily concentrate on F0 height and slope. I show that anticipatory planning affects the entire contour, as well as tone-segment alignment. I focus on four bimoraic forms, CVN, CV:N, CV_1V_2, and CV_1V_2N, each associated with F(alling) and R(ising) tones. These words were combined into all possible two-tone sequences (F+F, F+R, R+F, R+R), which speakers (3F, 1M) produced in a frame sentence that flanked the target sequence with mid tones.

The clearest dissimilatory effects occur in sequences of different tones (F+R and R+F). In these sequences, the entire contour of Target 1 is shifted toward the edge of the pitch range. F_1 starts higher ($p = 0.0002$) and has a higher peak ($p < 0.0001$) in F_1+R_2 than in F_1+F_2, while R_1 starts lower ($p = 0.01$) and has a lower valley ($p < 0.0001$) in R_1+F_2 than in R_1+R_2. The elbows are not more extreme solely by virtue of the full contour being shifted; the pitch excursions are also larger in F+R and R+F ($p < 0.0001$ and $p = 0.01$, respectively; see Figure 1).

This pitch dissimilation also delays the elbows in Target 1 relative to the onset of the word. The F_1 peak occurs later in F_1+R_2 than in F_1+F_2 ($p < 0.0001$), and the R_1 valley occurs later in R_1+F_2 than in R_1+R_2 ($p = 0.003$). This delay in timing is not a case of tonal (de)crowding: F elbows systematically occur earlier than R elbows (55% through the word for F, and 68% for R; $p < 0.0001$), but the elbow delay would have to be exclusively present in F_1+R_2, where the late R_2 elbow allows F_1 to shift rightward. However, the R_1 valley is also delayed when followed by the early-peak F_2, which is the opposite of tonal crowding.

Despite these shifts in contour timing, the segments remain constant, contrary to the segment-driven hypothesis. In none of the anticipatory shifts does the duration of the first mora shift with the elbow (non-significant interaction: $p = 0.94$). Even for the simplest of cases, early-elbow F vs. late-elbow R, the duration of the first mora is the same irrespective of tone ($p = 0.55$, see Figure 2). These results support the hypothesis that tones drive their own timing and are affected more by adjacent tones than by their segmental carriers.

Figures: Based on 255 trials from a representative participant (63 F+F, 64 F+R, 64 R+F, 64 R+R)

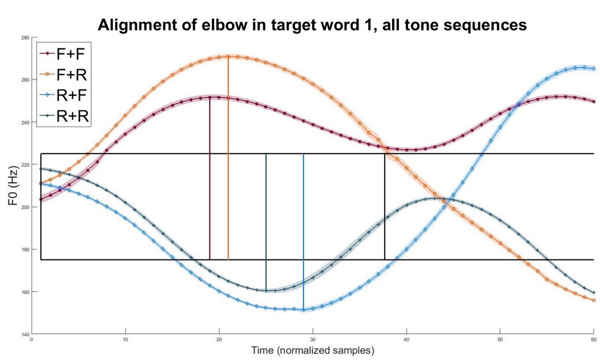

Figure 1: The initial portion of all tone sequences, with the elbow of target word 1 marked.

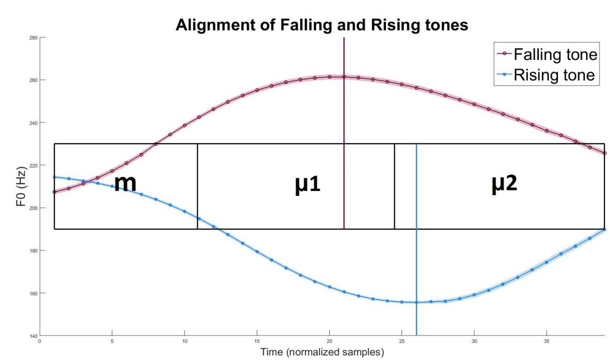

Figure 2: F0 contours for F and R, with elbows marked relative to mora edges.