

## The independence of the timing of tones and tone-bearing units: a study on Thai

Robin Karlin, Cornell University — [rpk83@cornell.edu](mailto:rpk83@cornell.edu)

**Introduction:** Thai words are typically monosyllabic, but lexical tones are not uniformly distributed. The distribution of tones is elegantly captured by a moraic tone-bearing unit (TBU): the two contour tones, Falling (HL) and Rising (LH), are restricted to words with two sonorant moras (e.g., /man/, /maa/), while the High (H) and Low (L) tones may also occur on forms with one sonorant mora (e.g., /mat/) (Morén & Zsiga, 2006). The contour tones are true contours, with pitch extrema (“elbows”) located near the midpoint of the word; the “falling” tone first rises and then falls, while the “rising” tone falls and then rises. Morén and Zsiga (2006) claim that this shape is the result of association to the right edge of the mora (see Fig. 1). Their right-edge analysis holds that lexical tone timing is driven by segmental timing, and predicts that tones will be invariably timed with the right edge of the mora.

I present results from an acoustic study in support of an alternate hypothesis: lexical tones generate internal timing independently of segments. Following Karlin 2014, this predicts that tones will be coordinated with each other more consistently than with segments. I test this hypothesis by manipulating environment and word shape, focusing on four bimoraic forms: CVN (/man/, /mun/), CVVN (/maan/, /muun/), CV<sub>1</sub>V<sub>2</sub> (/mia/, /mua/), and CV<sub>1</sub>V<sub>2</sub>N (/mian/, /muan/), each associated with F(alling, HL) and R(ising, LH) tones. These words were combined to form all possible two-word tonal sequences (F+F, F+R, R+F, and R+R), which speakers (3F, 1M) produced in a frame sentence (see Fig. 2).

**Results:** Elbow timing differed significantly across the two tones: the elbow occurs at approximately 55% through the word for the F tone, and at 68% for the R tone ( $p < 0.0001$ ; see Fig. 3). However, within a given word shape, the duration of the first mora is the same irrespective of tone ( $p = 0.55$ ). This results in a negative time lag between the right edge of the first mora and the tonal elbow for F tones, and a positive time lag for R tones (see Fig. 3).

Further evidence for the dissociation of mora edges and tone is provided by focusing on the diphthong stimuli. Both V<sub>1</sub> and V<sub>2</sub> are shorter in CV<sub>1</sub>V<sub>2</sub>N than in CV<sub>1</sub>V<sub>2</sub> ( $p < 0.0001$ ), but the elbows stay at 55% and 68% for the F and R tones, respectively. Thus, the right edge of the first mora is temporally more distant from the tonal elbows in CV<sub>1</sub>V<sub>2</sub>N than in CV<sub>1</sub>V<sub>2</sub> ( $p < 0.0001$ ).

It is also not the case that the right edge of the second mora is aligned with the end of the contour tone. This effect is visible on target word 1 in F+F and R+R sequences. In both cases, the tone of target word 1 continues on its trajectory beyond the limits of the word it is associated with. Thus, the valley between two falling tones and the peak between two rising tones both occur within the onset of target word 2 (see Fig. 4).

Finally, the timing of the tonal elbow in target word 1 is affected by the tone of target word 2. When the two tones are the same (i.e., F+F or R+R), the elbow of target word 1 is earlier than when the two tones are different (i.e., F+R or R+F); the interaction is significant ( $p < 0.0001$ ; see Fig. 5). However, the duration of the first mora does not change in parallel with the shifts in elbow timing, continuing in the spirit of tone-mora dissociation (non-significant interaction:  $p = 0.94$ ). That is, the tonal elbows move independently of segmental timing.

These results show that neither the tonal elbow nor the right edge of the tone is bounded by the TBU in Thai, despite the role the TBU plays in phonological distribution. Furthermore, there is an interaction between adjacent tones that is not predicted by straightforward TBU alignment. These results support the hypothesis that tones are strongly coordinated with each other, a relationship which may override segmental structures.

## References

- Karlin, R. (2014). The articulatory TBU: Gestural coordination of lexical tone in Thai. *Cornell Working Papers in Phonetics and Phonology*.
- Morén, B., & Zsiga, E. (2006). The lexical and post-lexical phonology of Thai tones. *Natural Language & Linguistic Theory*, 24(1), 113–178.

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

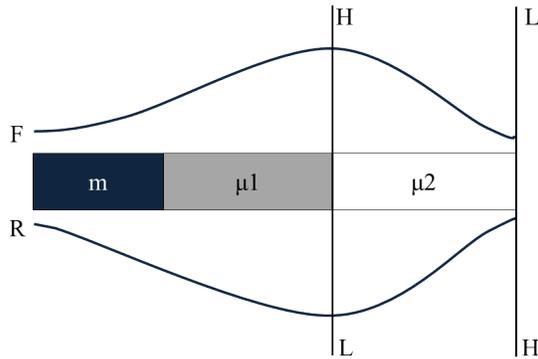


Figure 1: A schematic of predicted F and R tone shapes, according to the right edge hypothesis (Morén & Zsiga, 2006).

Kin	Targ. word 1	Targ. word 2	Adv
<i>Ms.</i>	<i>name</i>	<i>verbs</i>	<i>well</i>
khun	mâan-F	mûa-F	diidii
naang	mîa-F	mũun-R	diidii
naang	mñan-R	mûn-F	diidii
khun	măn-R	mũan-R	diidii

Figure 2: Example phrases used in the present study. The sequence of target words is flanked by mid tones. Kinship terms are assigned based on the vowel height of target word 1—*khun* before low vowels, and *naang* before high vowels.

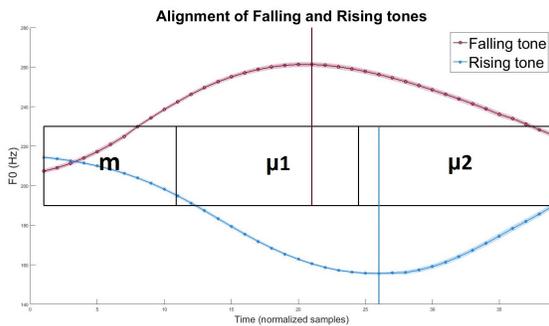


Figure 3: F0 contours for F and R, with elbows marked relative to mora edges.

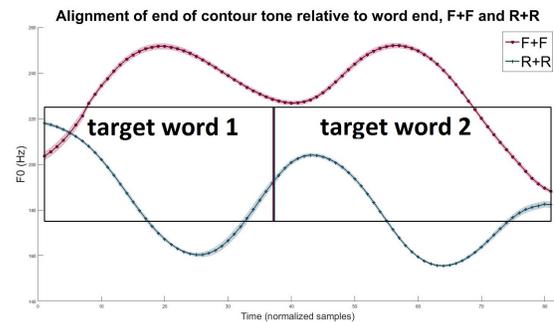


Figure 4: Trajectories for F+F and R+R, where the tone of target word 1 surpasses the word edge.

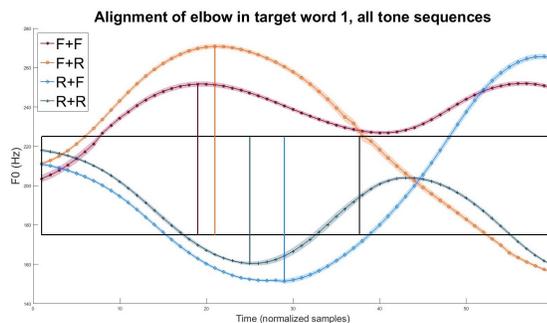


Figure 5: The initial portion of all tone sequences, with the elbow of target word 1 marked.