Realizations and Representations of Thai Tones in Monomoraic Syllables

Rattima Nitisaroj

Department of Linguistics
Georgetown University, Washington D.C., USA
rn29@georgetown.edu

Abstract

This study investigates the phonetic realizations of Thai tones in monomoraic syllables produced with three different speeds. The results do not support previous claim that tones in monomoraic syllables are neutralized to mid, and lend support to a phonological representation in which the mora is associated to either H or no tone at all. In normal speech, a /H/ vs. /∅/ opposition is found. In slow speech, a /H/ vs. /L/ distinction is found. Past research used the contrastive tonal status observed in slow speech as evidence to posit the underlying high and low tones for monomoraic syllables. In the present study, it is argued that the moraic association of the phonological low tone takes place as a result of glottal stop insertion.

Index Terms: Thai tone, monomoraic/linker syllable, speech rate, phonetics, phonology

1. Introduction

In the past, research on Thai tones has concentrated on the five tones in citation forms and stressed syllables (see, for example, [1]-[5]). To make a (near-)minimal quintet for tonal comparison, the syllables under investigation are always bimoraic, i.e., of the type CV:, CVC, or CV:C. In addition to bimoraic syllables, monomoraic syllables also exist in the language. They are of the type CV and have more restrictions with respect to the number of tones that are found on them. "Linker syllable" is the term [6] uses to refer to a syllable with the vowel /a/, which is usually realized as [ə]. In slow speech, the linker syllable ends with a glottal stop and carries a tone. In normal speech, the linker syllable loses the final glottal stop and its tone is adjusted to the environment. Lexically, monomoraic syllables have been argued to have a /H /vs. /L/ distinction, and the tones are neutralized to /∅/ at the surface level in normal speech [7]. Phonetically, previous studies agree that in normal speech, linker syllables are realized with a mid level pitch, and in slow speech, realizations of high and low tones are found [8, 9].

In the present study, an acoustic experiment is conducted to find out whether or not tones in monomoraic syllables can be distinguished from mid tones. Since it has been claimed that at the slower rate, there exists a distinction between high and low tones, speech rate also constitutes an independent factor in the experiment. From the acoustic results, it will be argued that that monomoraic syllables are not always phonetically realized with a mid level pitch and tonal specification is needed for a certain case. In addition, it will be shown that the realizations of low tones in slow speech may not be used as evidence to posit an underlying low tone because in this case, low tones show up as a result of glottal stop epenthesis.

The next section provides some details of the experiment. The acoustic results can be found in Section 3. Given the experimental results, Section 4 discusses the issues of phonetic realizations and phonological representations of tones in monomoraic syllables. Conclusions are provided in the same section.

2. Experiment

The present study is part of a larger study on the effects of stress and speaking rate on duration and tone in Thai [10]. The descriptions below address the part of the experiment that is relevant to the investigation of tones in monomoraic syllables.

2.1. Materials

Three disyllabic words were used: /mána:w/ 'lemon', /pàtʰá/ 'fight', and /na:ra:j/ 'Narayana'. The target syllables constituted the initial syllables in these words. The high and low tones placed on the monomoraic syllables reflected traditional descriptions. The target syllable in the last token contained a long vowel and carried a mid tone. As a result, the tones for comparison included the high tone on a linker syllable, the low tone on a linker syllable, and the mid tone on a long open syllable. To minimize carry-over and declination effects, all tokens were embedded in the initial position of appropriate sentence frames. Each of the sentences was then preceded by another sentence providing context.

2.2. Subjects

Nine female speakers of Thai participated in the study. All subjects were born and raised in Bangkok, Thailand until finishing at least the Bachelor's degree. Three of the subjects went abroad for a couple of years to earn their Master's. At the time of recording, all of the subjects worked in Bangkok. The average age of the subjects is 29.1 years old (SD = 2.67).

2.3. Recording Procedure

The target and context sentences were typed in Thai script on index cards. The subjects were asked to read the cards at three speech rates. The "normal" rate means a relaxed and comfortable tempo. The "slow" rate refers to the slowest tempo that the subjects could produce without inserting a pause between words within the same sentence. The "fast" rate is defined as the fastest tempo possible without making errors or losing naturalness. During the recording session, the subjects were first asked to read all the sentences at the normal rate. Then, they read at the slow and the fast rate. Some filler sentences were also added between the slow and fast rates. The subjects were asked to read these filler sentences at the normal

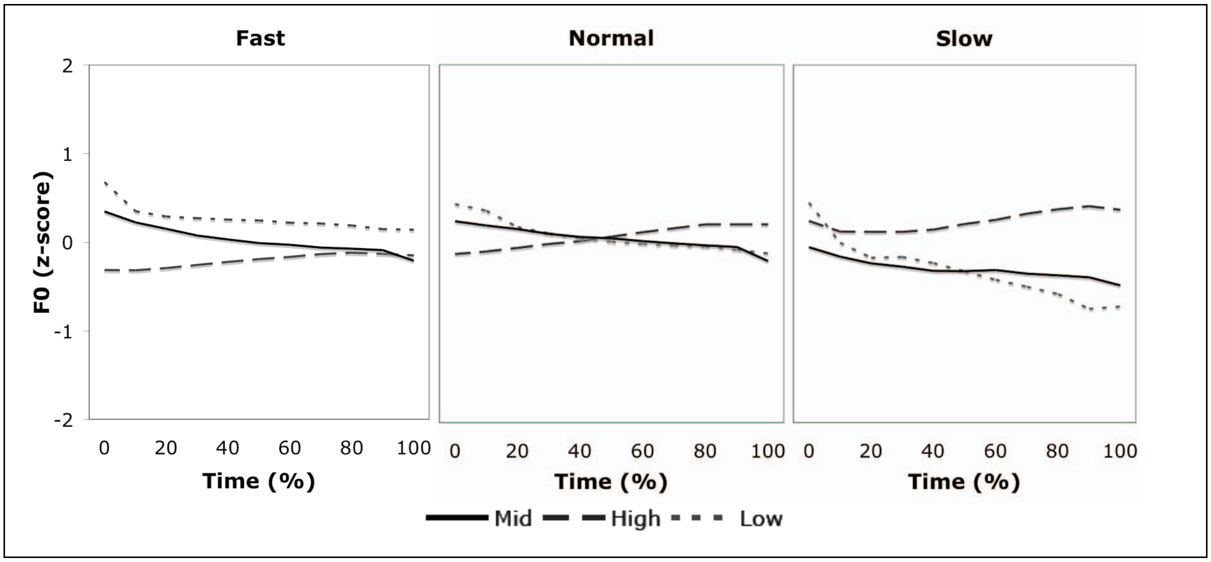

Figure 1. *Average F0-time-normalized contours of all tones at various rates.*

rate. These intervening filler sentences were put in to remind the subjects what the "normal" rate was and helped them to produce the "fast" rate as actually faster than the "normal" one. The materials were recorded directly onto a PC at a 22.05 kHz sampling rate using the Praat program [11].

2.4. Measurements

As part of a larger study, the rate manipulation was successful as confirmed by the mean rhyme durations for the different rates: 126, 163, and 205 msec for the fast, normal, and slow rate respectively. F0 was then extracted via an autocorrelation algorithm in the Praat program, using a 30 ms Hanning window with a step size of 5 ms. To be able to compare F0 configurations across syllable and tone types, F0 data were time-normalized by taking measurements at 11 locations throughout the rhyme of each target syllable. Measurements started from the beginning of the rhyme (0%) and moved every 10% of the rhyme duration until the end.

2.5. Data Normalization

Pitch range differences among the subjects were normalized by transforming Hertz values to a z-score scale [12, 13] using the following formula:

$$F0_{norm} = (F0_i - \bar{F}0)/s \quad (1)$$

where s is one standard deviation about the mean F0 ($\bar{F}0$)

Normalization parameters, i.e. mean and standard deviation, were calculated from raw F0 values of all tokens for a subject. The normalized F0 values thus express individual F0 values as the number of standard deviations above or below a subject's mean.

3. Results

Figure 1 presents the time-normalized contours of the tones extracted from the rhyme portions of the target syllables produced at the fast, normal, and slow rates. Tones in CV syllables are examined and compared with mid tones with respect to the six properties of average F0 height, highest F0 value, lowest F0 value, excursion size, highest F0 location, and lowest F0 location. A series of two-way repeated measure ANOVAs (rate x tone) were conducted to evaluate the extent to which these six properties varied given different rates and tones. The results are reported below.

3.1. Average F0 Height

Average F0 height defines the value averaged over the 11 measurement points of a syllable. A two-way repeated measure ANOVA indicated no main effect of either rate or tone on F0 heights. However, the interaction between the two factors was significant, $F(4, 32) = 4.18$, $p < .05$. At the fast and normal rates, there was no significant different between any tones. At the slow rate, mid and low tones were significantly lower in F0 height than high tones ($p < .05$).

3.2. Highest F0 Value

The repeated measure ANOVA conducted on highest F0 values revealed a main effect of tone, $F(2, 16) = 5.55$, $p < .05$, and a significant interaction between rate and tone, $F(4, 32) = 2.82$, $p < .05$. As indicated by post hoc Bonferroni tests, low tones had significantly higher F0 maxima than others ($p < .05$). This may be partly due to a consonant effect. The target syllable with the low tone begins with /p/ whereas others start with a nasal consonant. When each rate was considered, all differences between the tones were not significant except the difference between high and low tones at the fast rate ($p < .05$).

3.3. Lowest F0 Value

As indicated by an ANOVA, rate and tone interacted to produce a significant effect on lowest F0 values, $F(4, 32) = 5.05$, $p < .05$). There existed no main effect of individual factors. Within each speech rate, no significant difference was found except between high tones and others at the slow rate (*p*s < .05).

3.4. Excursion Size

Excursion sizes were calculated by taking the difference between the highest and the lowest F0 for each syllable. With respect to excursion size, an ANOVA indicated that significant effects came from tone, $F(2, 16) = 7.09$, $p < .05$, and the interaction between rate and tone, $F(4, 32) = 2.93$, $p < .05$. Post hoc Bonferroni tests revealed that in overall low tones had a significantly larger excursion size than high tones ($p < .05$). Regarding tonal contrast within rate, significant differences were found only at the slow rate. Low tones were significantly larger in excursion size than mid and high tones (*p*s < .05). This may be partly due to the onset consonant effect. However, low tones at the slow rate were still lower at the end.

3.5. Highest F0 Location

The results from the two-way repeated measure ANOVA conducted on highest F0 locations showed that tone was the only factor that produced a significant effect, $F(2, 16) = 10.43$, $p < .05$. Highest F0 values were found significantly later in syllables bearing a high tone than those carrying either a mid or a low tone (*p*s < .05).

3.6. Lowest F0 Location

Regarding lowest F0 location, a two-way ANOVA with rate and tone as independent variables displayed a main effect of tone, $F(2, 16) = 12.52$, $p < .05$. Post hoc Bonferroni tests demonstrated that mid tones were found with significantly later lowest F0 locations than high tones ($p < .001$).

The following table presents a summary of the acoustic properties that distinguishes between a pair of tones in a statistically significant manner.

Table 1. *Acoustic properties significantly distinguishing tones at different rates.*

		High	Low
Fast	Mid	Highest Location Lowest Location	---
	High		Highest Value Highest Location
Normal	Mid	Highest Location Lowest Location	---
	High		Highest Location
Slow	Mid	Average Height Lowest Value Highest Location Lowest Location	Excursion Size
	High		Average Height Lowest Value Excursion Size Highest Location

4. Discussion and Conclusions

This study has investigated six acoustic properties of Thai tones in monomoraic syllables uttered at the fast, normal, and slow rates. The tones were compared to mid tones in long open syllables. The results showed that mid and high tones can be distinguished at every rate, low and high tones can be distinguished at every rate, and mid and low tones can be distinguished only at the slow rate. Therefore, at the fast and normal rates, there is a two-way distinction: high and non-high. At the slow rate, the distinction is three-way: high, mid, and low. The results lead to the conclusion that in fast and normal speech, tones in monomoraic syllables demonstrate a distinction between high and non-high tones. In slow speech, the distinction becomes high vs. low. Previous claim and findings that high and low tones in "linker syllables" were completely neutralized to mid (see [8, 9]) is thus not supported by the current results.

Assuming the direct correspondence between phonetics and phonology, the following phonological representations are proposed.

- For fast and normal speech

```
High        Non-high
 H
 |
 μ           μ
 |           |
 V           V
```

- For slow speech

```
High        Low
 H           L
/ \         / \
μ  μ        μ  μ
|  |        |  |
V  ʔ        V  ʔ
```

At the fast and normal rates, the high vs. non-high distinction is reflected in the phonology as the presence or absence of the tonal target H associated to the mora. Assuming Stratal OT [14], lexical and post-lexical phenomena are differentiated. [5] provide a lexical grammar that allows all and only five lexical tones to surface in Thai. The present study assumes their constraint ranking and the lexical outputs generated by the ranking. These lexical outputs then serve as the inputs to the post-lexical component, which accounts for the realizations of tones in stressed and unstressed positions. Since monomoraic syllables are unstressed in Thai, the following part addresses the tonal realizations in unstressed syllables.

Given [5]'s ranking where *[TT]μ (a markedness constraint prohibiting any association of two tones to a single mora) dominates MAX[T] (a faithfulness constraint requiring the output to maintain the tone found in the input), the lexical output, i.e., the input to the post-lexical grammar, cannot have more than one tone associated to a mora. Therefore, for unstressed positions, the input forms are limited to [H], [L], and [∅]. When the input has H associated to the mora, the tone surfaces in the output. This arises from the domination of the faithfulness constraint MAX[H] (which requires the presence of the high tone in the output if the input has one) over the markedness constraint *[T] (which prohibits any association of tones) [15, 16]. In the case where the input mora is linked to L, it is better for the output to not have a tone than to maintain the tonal correspondence with the input. To allow for the lack of low tones in the output, the markedness constraint *[T] outranks the faithfulness constraint MAX[L]. As the constraint banning tones plays an important role in the system, no tone will be inserted if the input does not have one. Therefore, the proposed constraint ranking for tone associations in monomoraic/unstressed syllables in Thai is MAX[H] >> *[T] >> MAX[L]. By the given ranking of the markedness and faithfulness constraints, H is preserved while L is lost.

At the slow rate, the syllables become bimoraic or stressed, and both high and low tones are found. Crucially, here a glottal stop is inserted to CV syllables so the syllables become bimoraic. Tone associations thus pattern in the same way as those in obstruent-final syllables. In this case, the low tone emerges as a result of glottal stop epenthesis. Working within the OT framework, [5] propose the constraint C.G.Coda-> L, which requires the association of glottal codas with low tones. The constraint is supported by articulatory, acoustic, and perceptual evidence. By ranking C.G.Coda-> L over *[L], which prohibits low tones, an obstruent-ending syllable surfaces with a low tone. Since high tones are also found in CVO syllables, MAX[H], which bans deletion of high tones, must dominate C.G.Coda-> L. See [5] for a detailed OT analysis of tones in syllables with an obstruent coda.

In summary, at the normal rate, tones in monomoraic syllables are not completely neutralized. Assuming a constraint-based analysis, the language-specific constraint rankings will allow the high tones to surface if the input has it. Given the input with a low tone or no tone, the constraint interaction yields the output with no associated tone. The phonological /H/ vs. /∅/ opposition is the most common pattern for a two-level tone system [17]. The results in this study show that Thai follows this widely observed characteristics, instead of having a less frequent /H/ vs. /L/ opposition. In addition, this study argues that the phonetic manifestation of low tones in slow speech is not the realization of the underlying low tone, but arises from glottal stop epenthesis and the requirement that glottal codas must be associated with low tones.

5. Acknowledgements

I wish to thank Dr. Elizabeth Zsiga for her valuable advice and comments.

6. References

[1] Abramson, A. S., *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments,* Bloomington: Indiana Research Center in Anthropology, Folklore, and Linguistics, 1962.

[2] Gandour, J., Potisuk, S., Ponglorpisit, S., and Dechongkit, S., "Inter- and intraspeaker variability in fundamental frequency of Thai tones," *Speech Communication,* vol. 10, pp. 355-372, 1991.

[3] Potisuk, S., Gandour, J., and Harper, M. P., "Contextual variations in trisyllabic sequences of Thai tones," *Phonetica,* vol. 54, pp. 22-42, 1997.

[4] Tingsabadh, M. R. K., and Deeprasert, D., "Tones in Standard Thai connected speech," in *Southeast Asian Linguistic Studies in Honour of Vichin Panupong,* ed. A. Abramson, Bangkok: Chulalongkorn UP, 1997, pp. 297-307.

[5] Morén, B., and Zsiga, E., "The lexical and post-lexical phonology of Thai tones," *Natural Language and Linguistic Theory,* vol. 24, pp. 113-178, 2006.

[6] Bee, P., "Restricted phonology in certain Thai linker-syllables," in *Studies in Tai Linguistics in Honor of William J. Gedney,* ed. J. G. Harris, and J. R. Chamberlain, Bangkok: Central Institute of English Language, 1975, pp. 17-32.

[7] Bennett, J. F., "Metrical foot structure in Thai and Kayah Li: Optimality-Theoretic studies in the prosody of two Southeast Asian languages." PhD diss., University of Illinois at Urbana-Champaign, 1995.

[8] Hiranburana, S., "Changes in the pitch contours of unaccented syllables in spoken Thai," in *Tai Phonetics and Phonology,* ed. J. G. Harris, and R. B. Noss, Bangkok: Central Institute of English Language, 1972, pp. 23-27.

[9] Luksaneeyanawin, S., "Intonation in Thai." PhD diss., University of Edinburgh, 1983.

[10] Nitisaroj, R., "Effects of stress and speaking rate on duration and tone in Thai." PhD diss., Georgetown University, 2006.

[11] Boersma, P., and Weenink, D., *Praat: Doing phonetics by computer* (Version 4.3.01) [Computer program], 2005. Retrieved from http://www.praat.org/.

[12] Jassem, W., "Normalization of F0 curves," in *Auditory Analysis and Perception of Speech,* ed. G. Fant and M. Tatham, London: Academic Press, 1975, pp. 523-530.

[13] Rose, P., "Considerations in the normalization of the fundamental frequency of linguistic tone," *Speech Communication,* vol. 6, pp. 343-351, 1987.

[14] Kiparsky, P., "Opacity and cyclicity." *Linguistic Inquiry,* vol. 17, pp. 351-365, 2000.

[15] Bickmore, L., "Bantu tone spreading and displacement as alignment and minimal misalignment. Ms., University of Albany, 1996.

[16] Yip, M., *Tone,* Cambridge: Cambridge UP, 2002.

[17] Hyman, L., "Privative tone in Bantu." Paper presented at Symposium on Tone, ILCAA, Tokyo, 2000.