On the acoustic reality of "toneless" morphemes in Thai: Preliminary investigations from sequence-internal final particles

Ratanon Jiamsundutsadee✉
Department of Linguistics, Faculty of Arts, Chulalongkorn University

(Received: 2nd October 2022; Revised: 2nd December 2022; Accepted: 19th December 2022)

Abstract

This study aims to examine and compare the F0 behavior of tonal final particles (FPs) and that of toneless FPs in Thai when occurring inside the intonational phrase based on naturalistic speech data. The results reveal that tonal FPs have significantly greater average rate of change values than toneless FPs, and that toneless FPs show more substantial variability in pitch range. Interestingly, cases such as Tone 2 and L% FPs do not pattern with the general trend in terms of rates of change; regardless, difference in pitch still functions as a cue to their distinction. On the whole, these findings support structural analyses which make a representational distinction between the two types of FPs in Thai.

Keywords: tone, tonal underspecification, final particles, Thai

Introduction

In broadest terms, tone languages are defined as those whose pitch can distinguish meaning at the word level (Yip, 2002). One assumption arising from such a definition is that pitch must be specified in the lexical representation of "at least some morphemes" (Hyman, 2001). Accordingly, the notion that certain morphemes in a given tone language may lack underlying tonal specification has been widely discussed in the tonal and intonational literature (e.g. Yip, 1980; Odden, 1982; McPherson, 2011) in parallel with segmental underspecification (e.g. Keating, 1988; Cohn, 1993; Choi, 1995). Following the autosegmental-metrical (AM) approach (see Pierrehumbert, 1980), predictions could then be made in which these tonally underspecified morphemes would exhibit surface pitch contours depending on the neighboring context, transitioning between the respective surrounding tonal specifications. Typologically, toneless morphemes are often grammatical, with tonal specification reserved primarily for lexical morphemes (Hyman, 2001). Grammatical morphemes include postpositions (Hyman, 2001), discourse markers (Greif, 2010), and final particles (Yip, 1980; Law, 1990).

Thai[1] serves as a pertinent case study where the phenomenon of neutral(ized) tone has previously been independently reported in unstressed syllables (Kallayanamit, 2004; Luksaneeyanawin, 1983) and grammatical morphemes (Bee, 1975; Henderson, 1949). Even more intriguing are final particles (FP): it is generally agreed upon that their surface variability stems

✉ ratanonj@outlook.com

[1] "Thai" in the linguistics literature typically refers to "Bangkok Thai", "Central Thai", or less likely "Standard Thai". In this study, "Thai" is used to loosely refer to "Bangkok Thai" and "Central Thai", as the factor of naturalness considered during the design of the stimulus is based on the language variety spoken in the Bangkok Metropolitan Region, in which all participants either were born or raised.

from the lack of an underlying tonal specification (Henderson, 1949; Luksaneeyanawin, 1983). One of the most recent accounts by Pittayaporn (2014) proposes the intonational boundary tones H% and L% and categorizes FPs into inherently tonal and toneless. Toneless FPs are said to associate with boundary tones. On the other hand, lexical tones as specified in tonal FPs override boundary tones, prohibiting them from appearing in the surface representation. A crucial trait of these boundary tones is that only one of them is allowed irrespective of the number of FPs, and that they occur at the right edge of the intonational phrase. Consequently, toneless FPs are not linked to a boundary tone when appearing inside the intonational phrase, and their surface realizations are simply a transition from tone targets occurring before and after the FP, compared to the tonal stability of tonal FPs regardless of their location. Reflecting the phonological nature of each type of FP, phonetic behavior such as that of fundamental frequency (F0) is thus a potential key to distinguishing between the two FP categories embedded within the intonational phrase.

Gathering naturalistic speech data to address problems and inquiries in phonology, such as the one currently at hand, requires attentiveness to the experimental design process. The decision to utilize a situational dialogue reading task stems from the current study's attempt to overcome the limitations of various elicitation techniques, comprising controlled stimuli in a natural environment.

This study aims to examine and compare the F0 behavior of tonal FPs with that of toneless FPs in Thai. Specifically, the matter of investigation is whether tonal and toneless FPs in Thai exhibit similar F0 behavior when occurring inside the intonational phrase. In line with phonetic and phonological accounts regarding tonal underspecification and the tonal-intonational characteristics of Thai, it is hypothesized that, as tonal FPs are specified with underlying pitch targets, they will have greater average rate of change than phonologically underspecified toneless FPs in Thai.

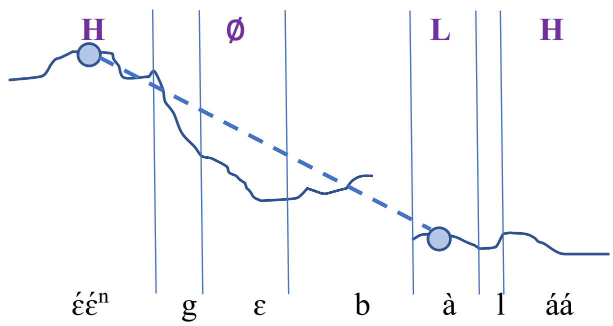
Tonelessness and tonal underspecification

Tone is a notorious concept that has generated much controversy among phonologists, phoneticians, and linguists of other subfields. Adopting Hyman's (2001, as cited in Yip, 2002) definition, tone languages are those "in which an indication of pitch enters into the lexical realization of at least some morphemes." Accordingly, if the specification of pitch only applies to "(at least) some morphemes" — not all — it is therefore possible that a tone-bearing unit (TBU) in a given tone language, be it the vowel, mora, or syllable, is "toneless". Such toneless TBUs are not supplied with tonal specification at the phonological level, as attested in several tone languages. Notably, in Hyman's (2001) treatment of the tonal system of Bantu languages, it is pointed out that one crucial trait is that the surface realization of pitch can vary by contextual conditions because pitch is determined solely from the interpolation between the surrounding pitch targets.

An influential approach to integrating the notion of tone into phonology is the autosegmental-metrical (AM) theory. In brief, phonological abstractions of pitch comprise autosegments, namely H(igh) and L(ow) tones. These autosegments are then structured sequentially, one after another and never overlapping, to give a string of local events which correspond to specific parts in the segmental level. Therefore, in between these events where no autosegments are present, i.e. where segmental units are phonologically unspecified with respect to pitch, the surface pitch contour is *transitory*, dependent on the prior and following events or specifications of pitch (Ladd, 2008; Arvaniti & Fletcher, 2020). For example, the pitch track of a Tommo So[2] utterance in Figure 1 illustrates how the F0 smoothly interpolates from one target to another, with the toneless syllable's contour predictable from the surrounding context.

Figure 1

Pitch track of [έέ^n=gɛ bàláá] 'she swept up the ashes'
(reproduced and adapted from McPherson, 2011)



Despite the AM framework being historically developed from autosegmental analyses of African tone languages (Lionnet & Hyman, 2018), it has since been widely applied to languages in Asia, whether tonal (Cantonese – Wong et al., 2005; Mandarin – Frazier, 2014; Vietnamese – Ha, 2012) or non-tonal (Jeddah Arabic – Moussa, 2019; Tamil – Keane, 2006, 2012). Thai is no exception, as can be seen in Morén & Zsiga's (2006) interpretation of Thai tonal system with mora as the TBU.

Phonetic characteristics of Thai tones

Five tones are lexically contrastive in Thai. Traditionally, linguists have dubbed them as "Mid", "High", Low", "Falling", and "Rising"; however, so as to avoid confusion with the respective phonetic realizations of each tone, numeric labels will instead be used in this study (see Table 1).

[2] Tommo So is a Niger-Congo language spoken in Mali (McPherson, 2013).

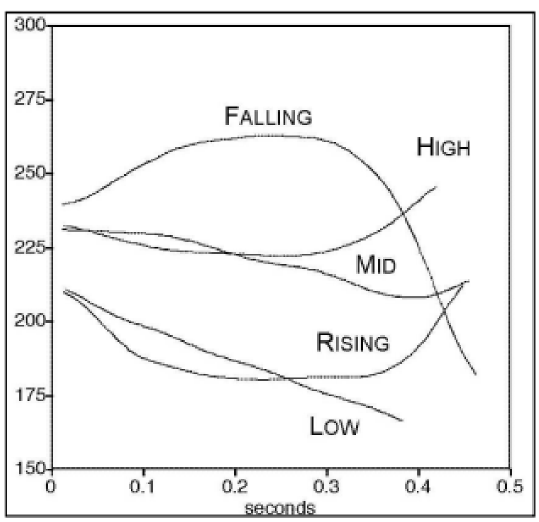
Table 1
Thai tones with examples

Tone	Thai Orthography	IPA Transcription	Gloss
Tone 1 (Mid)	ฟา	/faa/	‘fa’ (musical note)
Tone 2 (Low)	ฝ่า	/fàa/	‘to undergo’
Tone 3 (Falling)	ฝ้า	/fâa/	‘ceiling; haze’
Tone 4 (High)	ฟ้า	/fáa/	‘sky; blue’
Tone 5 (Rising)	ฝา	/fǎa/	‘lid, cover’

One of the more recent prominent works on Thai tones is that of Morén and Zsiga (2006). Albeit controversial in proposing mora as the TBU, as opposed to the more generally agreed upon syllable, their study presents what Thai tones in isolation looked like in the early 21st century, as depicted in Figure 2

Figure 2

Pitch track of Thai tones in citation form (in Hz) averaged across speakers (from Morén & Zsiga, 2006, as quoted in Zsiga, 2007, in Thepboriruk, 2009: 3)

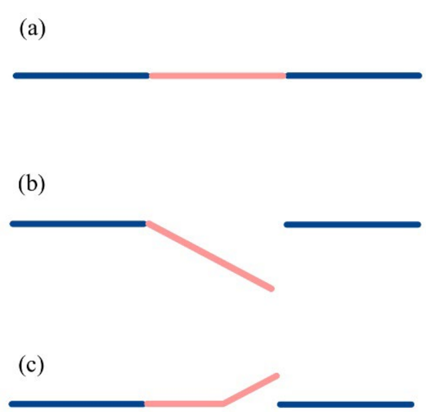


Looking specifically at Tone 2 (labelled as LOW) and Tone 4 (labelled as HIGH), as the target words in this study carry these two tones, Tone 2 gradually falls from slightly lower than mid range to low. Meanwhile, Tone 4 plateaus at mid range in the first half, then rises slightly in the latter.

In terms of connected speech, Thai lexical tones in stressed syllables exhibit no substantial change (namely tone sandhi and neutralization): tonal contrast is fully preserved. One major pertaining concept, which greatly affects the predictions of this study, is tonal coarticulation. Like segmental coarticulation, the coarticulation of tones refers not to how they are realized individually, one after another in discrete blocks, but to how they vary according to adjacent or nearby neighbors (Hardcastle & Hewlett, 2006). Gandour and Potisuk (1994) conducted a seminal investigation on tonal coarticulation in Thai, where they tested the 25 possible sequences of two consecutive tones realized through monosyllabic words in a carrier sentence. Surprisingly, they discovered three crucial properties: 1) Thai tones were more heavily influenced by perseverative than by anticipatory coarticulation; 2) coarticulation affected pitch height but not slope; and, in a related manner, 3) the anticipatory effects were mostly dissimilatory, instead of assimilatory. Highly pertinent to this study is the second finding, in which the contour of each tone is preserved, with only the height of the tonal onset affected by the preceding tone. In other words, regardless of their bordering context, Tone 2 will always be gradually falling, and Tone 4 will always be mid level with a final rise. Even further, if tonal syllables are embedded between identical neighboring tones, such as a [T1 ___ T1] environment, their respective contour will be distinct from that of a toneless syllable. Following the linear interpolation model discussed in 1.1, toneless syllables are predicted to have a more or less steady, transitory plateau from the first T1 syllable to the other. The relevant predictions are schematically illustrated in Figure 3.

Figure 3

Schematic representation of (a) toneless, (b) Tone 2, and (c) Tone 4 syllables (in pink) embedded in a [T1 ___ T1] environment (in blue)[3]



[3] Owing to the schematic nature of this representation, it does not faithfully reflect the fine (yet distinct) characteristics of each contour, such as the slight fall of Tone 1, and it also ignores other relevant phenomena, such as prosodic-phrase-final terminal junctures and declination. Nevertheless, the main purpose of this illustration is to visually motivate the measurement of rate of change in distinguishing between toneless and tonal syllables, as will be discussed further.

The consequence is that tonal syllables exhibit greater movement than their toneless counterparts. By movement, I am referring to the *rate of change* between two points in time, as will be detailed in the Methods section.

Tonal and intonational accounts of FPs in Thai

If, according to Bolinger (1978) and Gussenhoven (2004), intonation is a “half-tamed savage” due to its paradoxically linguistic and paralinguistic nature, I aver that FPs, at least in Thai, are shapeshifting apparitions. Even in one of the most comprehensive treatments of Thai FPs written by Cooke (1989), it is admitted right at the very beginning that “one of the most baffling areas of the Thai language… is the sentence-particle system” (Cooke, 1989: 2). One problem has to do with the amount of variability both in terms of form and function, while the other pertains the inability to pinpoint the exact meaning of each FP (hence a shapeshifting apparition).

Traditionally, every morpheme in Thai is analyzed as being specified with tone; as a result, many linguists treat FPs as also having an underlying tonal specification (Noss, 1964; Cooke, 1989). However, as Pittayaporn (2014) points out, this view fails to account for the clear correlation between (what is analyzed as) Tones 4 or 5 and questions, as well as Tones 2 or 3 and statements, considering them as occurring by chance. Another approach is to analyze FPs as intonational: one example comes from Luksaneeyanawin’s (1998) chapter on the intonational system of Thai. It is propounded that “the prosodic complex postulated as word prosody [i.e. tone] in [Henderson (1949) and Chuenkongchoo (1956)] should have been postulated as sentence or phrase prosody [i.e. intonation], since the complex does not affect only the final particle but the whole utterance” (Luksaneeyanawin, 1998: 393).

The only study which adopts the AM model in breaking down the tonality of FPs is Pittayaporn’s (2014) work on the tone-intonation interaction as seen through FPs in Thai. The main argument is that FPs can either be tonally specified or underspecified in their lexical representation. Moreover, lexical tones exist together with intonational tones (in the form of intonational-phrase-final boundary tones H% and L%) in a linear sequence. While H% is taken to be signaling a hearer-oriented proposition (such as questions, vocatives, and other utterances which require a response), L% is a speaker-oriented one (such as statements, directive speech acts, and other utterances) which expresses non-involvement. Lexical tones and boundary tones must compete in order to surface in the post-lexical representation. As such, the intonational tones always lose to their lexical counterparts; in consequence, boundary tones can only surface on tonally underspecified or toneless FPs. Cases in point are provided in (1)-(4).

(1) ɾaw^{1} cʰɔɔp^{3} taw^{2} kʰa-L%
1SG like turtle FP
'I like turtles.' (/kʰa/ = formal, female speaking)

(2) $\text{naa}^{3}\text{ɾak}^{4}$ maj^{4} kʰa-H%
cute FP FP
'Are they cute?' (/maj^{4}/ = neutral interrogative; /kʰa/ = formal, female speaking)

(3) ɾaw^{1} cʰɔɔp^{3} taw^{2} kʰɾap^{4}
1SG like turtle FP
'I like turtles.' (/kʰɾáp/ = formal, male speaking)

(4) naa^{3}ɾak^{4} maj^{4} kʰɾap^{4}
cute FP FP
'Are they cute?' (/maj^{4}/ = neutral interrogative; /kʰɾáp/ = formal, male speaking)

Critical evidence of the distinct lexical representation between toneless and tonal FPs, which serves as the basis of this study, comes from a stacked FP sequence. Two predictions arise from the analysis: 1) boundary tones are separate phonological units from FPs themselves, and they are *associated* at the post-lexical level; 2) there can be only one tonal target at the intonational-phrase-final position, whether it be lexical or boundary tones. Therefore, if a toneless FP resides *within* the intonational phrase, the tonal neutrality becomes apparent in its surface realization, as exemplified in (5).

(5) naa^{3}ɾak^{4} na kʰa-H%
cute FP FP
'They are cute.' (/na/ = common ground; /kʰa/ = formal, female speaking)

Relating back to the phonetic predictability of tonelessness, its contour is accordingly transitory, interpolating from the preceding target to the following. Conversely, tonal FPs are tonally stable regardless of their position along the intonational phrase. On a final note, it is important to note that this distinction is lost when both FPs in question occur at the right edge of the intonational phrase, as they will always surface with either an inherent lexical tone or an associated boundary tone.

Eliciting natural speech data

Phonological inquiries have long consulted two main sources of information, namely primary data from a language informant and secondary data based on other primary texts of the same origin (Hyman, 2006). The calls from various linguists for testing phonological hypotheses based on instrumental methods during the 70s and 80s, central to which was the rise of the AM theory, provided the impetus for laboratory phonology. Elicitation is thus a crucial means of gathering empirical, natural data to answer intriguing questions in phonology.

According to Faitaki and Murphy (2020), oral language elicitation tasks can be categorized in a continuum from lowest to highest control (see Ambridge & Rowland, 2013). Spontaneous speech data gives researchers the least control; however, as the spontaneity of elicitation might suggest, one of the most detrimental trade-offs pertaining to the current study is that spontaneous speech is "bound to be uninformative for researchers who wish to investigate… infrequent linguistic structures [or those too subtle to be perceived by the researcher]." (Faitaki & Murphy, 2020: 363). On the other end of the spectrum are straightforward, more controlled tasks such as eliciting speech from a written prompt. Researchers would be in full control of the stimuli, albeit in citation form; that is, the speech data might be useful solely for inquiries with cues that laypeople are unaware of or unaffected by.

Turning to the current study, sequence-internal FPs might not be uncommon among informal everyday conversations; however, hoping for the appropriate distribution of toneless and tonal FPs within specific tonal contexts might be nothing less than futile. In contrast, using written prompts for controlled elicitation tasks might not be the most appropriate, especially when considering that orthographic effects have an influence on how speakers elicit a word. Specifically, tone markers, or lack thereof, cue the existence of a tonal target for literate speakers of Thai. Hence, they are inclined to sound out what Pittayaporn (2014) analyzes as toneless syllables as if they were fully specified for tone.

To combat the methodological tug of war between naturalness and control over target stimuli, I have chosen to use a situational dialogue reading task. By balancing out the disadvantages of both spontaneous and citation speech, promising results are expected. Tackling the rarity of occurrence of intonational-phrase-internal FPs, I produced dialogues where all the target FPs are embedded in carrier sentences with appropriate tonal contexts. As for naturalness, the dialogues 1) concern everyday situations (in this case, a dog or a cat goes missing, and another person needs help finding it) and 2) comprise two interlocutors, both roleplaying as close friends of each other. These two factors greatly help shape a natural condition for speech elicitation, as opposed to the more rigid laboratory setting.

Methods

2.1 Participants

Eight female native speakers of Thai (Speakers A-H), aged 21, participated in this study. They are all fourth-year undergraduates currently studying at the Faculty of Arts, Chulalongkorn University. Each participant either was born or grew up in the Bangkok Metropolitan Region, predominantly speaks Bangkok Thai in everyday circumstances, and is unable to speak any other given regional varieties of Thai. No participant reported any forms of speech-language impairment. Everyone received monetary compensation for their participation in this study. Taking into consideration the relative colloquiality of the dialogue used in the elicitation task (as will be outlined below in the Procedure section), young adults were chosen as participants for this experiment. Detailed information about each participant is provided in Appendix A.

2.2 Stimuli

In order to compare the F0 behavior of intonational-phrase-internal tonal and toneless FPs, the criteria for the selection of appropriate stimuli are as follows:

1) Their entire syllable must be sonorant so that they are able to fully bear pitch;

2) they are generally used in everyday situations by Thai speakers of demographic similarity to the participants; and

3) they can be followed by มึง /mɯŋ^1/ at the intonational-phrase-final position (which point will be motivated below).

Due to the lack of sufficient targets, FPs with glottal stop codas are also allowed. Table 2 shows the target words selected for this study, comprising 8 target FPs (four toneless FPs and four tonal FPs).

Table 2
Target words

FP	Thai Orthography	FP Category	Gloss (adapted from Pittayaporn, 2014)
na(-H%)[4]	นะ	toneless	common ground
na(-L%)	น่ะ		
wa(-H%)	วะ		non-restraint
wa(-L%)	ว่ะ		
maj^4	มั้ย	Tone 4	neutral interrogative
nɤʔ^4	เนอะ		solicitation of involvement
niʔ^2	หนิ[5]	Tone 2	relevance
nɛʔ^2	แน่ะ		highlight

Each target word is embedded in a variety of carrier utterances, consisting of one intonational phrase per utterance. The target word is located at the penultimate (second-to-last) position of the phrase, preceded by a CaaC^1 syllable whose coda is a sonorant inverse to the onset of the target word (i.e. Caaj^1 or Caaw^1 for FPs with nasal onsets, and Caam^1 or Caan^1 for FPs with approximant onsets). The final monosyllable word, which the target FP precedes, is always the informal second-person singular vocative มึง /mɯŋ^1/.

The reasoning behind the choice to stack the FP behind /mɯŋ^1/ is to prevent paralinguistic effects from shifting or exaggerating the pitch contour of intonational-phrase-final FPs. Moreover,

---

[4] It is crucial to note once again that such boundary tones as H% and L% are not a part of each FP's lexical representation; on the contrary, they are associated with the FPs *post-lexically*. For example, /na/ only has one underlying representation. When /na/ is linked to either two boundary tones, however, it can realize as [na-H%] or [na-L%], with this surface distinction being perceived as different by Thai speakers' intuition and reflected in the orthography (i.e. tone markers or lack thereof).

[5] T2 FPs and those associated with L% often have two or more orthographic variants, chiefly one with a tone marker and another with ห นำ. In consideration of orthographic consistency, all but one target stimulus is presented with a tonal marker. This is due to the unfortunate fact that <นิ่> for /niʔ^2/ is markedly unnatural.

the decision to not also use FPs (instead of the pronominal vocative) for this purpose is based on the fact that only some FPs can be stacked in a specific linear order: question FP - modality FP - status FP (Pittayaporn, 2014). A prototypical example would be [na kʰa-H%], [na kʰa-L%], and [na kʰráp]; unfortunately, all such possible final FPs that are productive enough to function as the controlling context for this experiment have obstruent onsets. Despite their perturbations mostly affecting the beginning of the following vowel, and not to mention the non-sonorant portions themselves giving out inaccurate pitch track readings, the entire section comprising the target FPs and their surrounding contexts are needed for further analysis[6]. Therefore, disarranged F0 values from perturbation effects is unsatisfactory, especially for quantitative analyses, resulting in the need to make sure that the target and the context are wholly sonorant and able to bear pitch. A major problem is thus proving that /mɯŋ^{1}/ truly resides at the right edge of the intonational phrase as in (6b) instead of forming its own prosodic constituent separate from the target FPs as in (6c).

(6)	a.	jaa^{2}	kʰɾiət^{3}	na	mɯŋ^{1}
		don't	be.stressed.out	FP	2SG.VOC
		'Don't stress about it!'			
	b.	[jaa^{2} kʰɾiət^{3} na mɯŋ^{1}]$_{\iota}$			
	c.	[jaa^{2} kʰɾiət^{3} na]$_{\iota}$ [mɯŋ^{1}]$_{\iota}$			

Regardless, the FP targets as observed in this study were realized in a very similar way to the case where the intonational-phrase-final position is occupied by a true FP. For example, instead of an expected glottal stop for a citation-style reading of /na/, due to orthographic effects, no glottal stops are present. I have therefore analyzed /mɯŋ^{1}/ as an FP at the right edge of the intonational phrase.

2.3 Procedure

Each recording session was conducted in a quiet room, with the participant seated in front of a Signo MP-701 cardioid condenser microphone 20 centimeters away from their mouth. The dialogue script for each trial was presented on a laptop screen written with Thai orthography, deliberately choosing the less standard variants (e.g. <ยังไง> instead of <อย่างไร>) in order to fit the interpersonal context of the dialogue. In each set of dialogue, the target FPs were embedded in their respective carrier sentences at random, scattered across two separate dialogues within the set, although the content of the dialogue must read naturally and immediately make sense to the participant. Each participant was instructed to follow along the dialogue script, taking turns with the researcher. They were also instructed to read aloud the provided lines in a casual manner at a normal pace without stopping mid-sentence, as if they were talking to a close friend[7]. Four repetitions of two sets (each containing two dialogues) of dialogues were recorded, with each

[6] Due to several limitations, such analyses taking into consideration the contextual F0 behavior are not included in this study.
[7] These instructions were explicitly given before each set and were also written above each of the dialogue scripts.

set recorded in a single session. Each recording was made via Audacity (version 3.0.2) at a sampling rate of 44,100 Hz. The full script is provided in Appendix B.

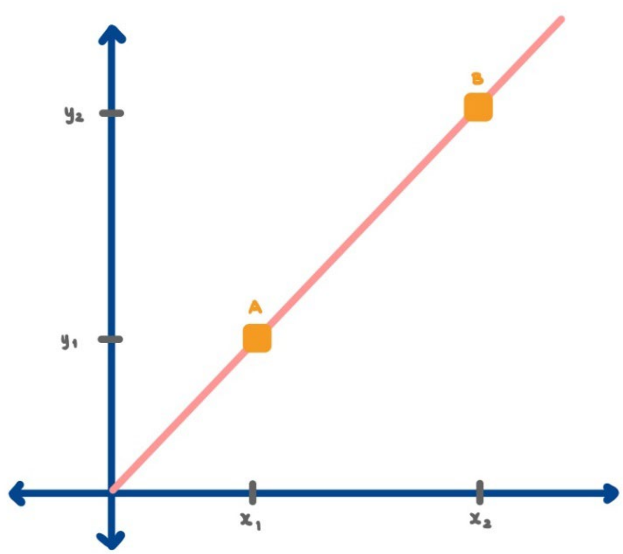
2.4 Acoustic measurements and analysis

Acoustic analysis was conducted on a total of 512 tokens (8 speakers × 8 target FPs × 2 sets × 4 repetitions), 88 tokens per participant. Recordings were exported from Audacity as separate WAV files for each set. Each token was then manually segmented and labeled in Praat. ProsodyPro (Xu, 2005-2021), an interactive Praat script, was slightly customized to automatically measure F0 values at 11 equidistant points from the onset to the offset of each contour. Target words in which 111 data points are missing or abnormal (such as sudden drops to near-zero marks), usually due to creakiness, were discarded. Further analyses were done based on the remaining 401 usable tokens.

There are several quantitative metrics used to calculate how "quickly" a phonetic parameter moves from one point to another, one of which is slope. For example, in order to measure how steep (therefore quick, and therefore indicative of more movement) the F0 from Point A to Point B in Figure 4 is, its slope can be calculated via the formula in (7).

(7) slope = $\frac{y_2 - y_1}{x_2 - x_1}$

Figure 4
An imaginary linear function



Additionally, in a linear function such as that depicted in Figure 4, the steeper the slope, the quicker the rate of change. Calculating the slope would be appropriate for Tone 2 with its more or less straight fall. However, for contours with inflection points such as Tone 4, the slope would be derived only from the onset and the offset of the contour, disregarding the rest of the F0 movements in between. A straight line and a concave or convex with starting and end points of the same coordinates would result in the same slope. Hence, a more suitable metric must be derived from finding the slopes between each time point within the contour, resulting in a set of slope values for the entire contour. These values are then averaged out so as to arrive at the average rapidity or rate of change for said contour.

The principal metric determined from each time-normalized contour is the average rapidity or rate of change (henceforth ROC). Also known as F0 attack (Ross et al., 1986), ROC calculates the slope for every time point of continuous voicing and averages each slope to acquire the contour's average ROC. Statistical analyses may then be based on this acoustic measure. The formula used to derive the average ROC for each token is shown in (8).

(8) $$\mu_{ROC}(\text{token}) = \frac{\frac{|F0_2 - F0_1|}{t_2 - t_1} + \frac{|F0_3 - F0_2|}{t_3 - t_2} + \dots + \frac{|F0_{n+1} - F0_n|}{t_{n+1} - t_n}}{n}$$

where n equals the number of slopes measured (in this case, 10 slopes calculated from 11 F0 data points), and the time points (t) are 0%, 10%, …, up until 100%.

Moreover, for graphical analyses which call for interspeaker comparisons, an F0 normalization method based on z-scoring (Zhang, 2018) was utilized in order to minimize the physiological and sociolinguistic differences between each participant. The formula is indicated in (9).

(9) $$F0_{z\text{-score}} = \frac{Hz_{raw} - m}{s}$$

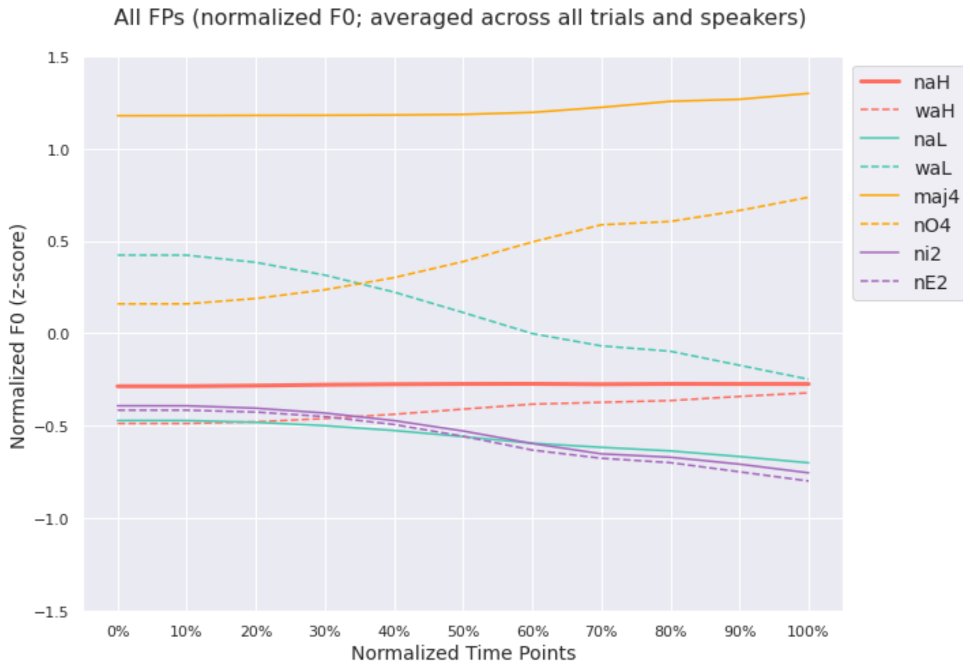
where Hz_{raw} is the raw, untransformed F0 value, and *m* and *s* are the mean value and the standard deviation calculated per speaker, respectively.

Statistical analyses, namely Welch's *t*-tests, were carried out on Microsoft Excel 2013. Pitch track visualizations were created through Matplotlib, a plotting and visualization library for Python, and its related library Seaborn, via Google Colaboratory cloud computing interface.

Results and Analysis

Figure 5 displays time- and F0-normalized contours for all target FPs, averaged across all speakers and their respective trials.

Figure 5
Time- and F0-normalized FPs, averaged across all speakers and trials

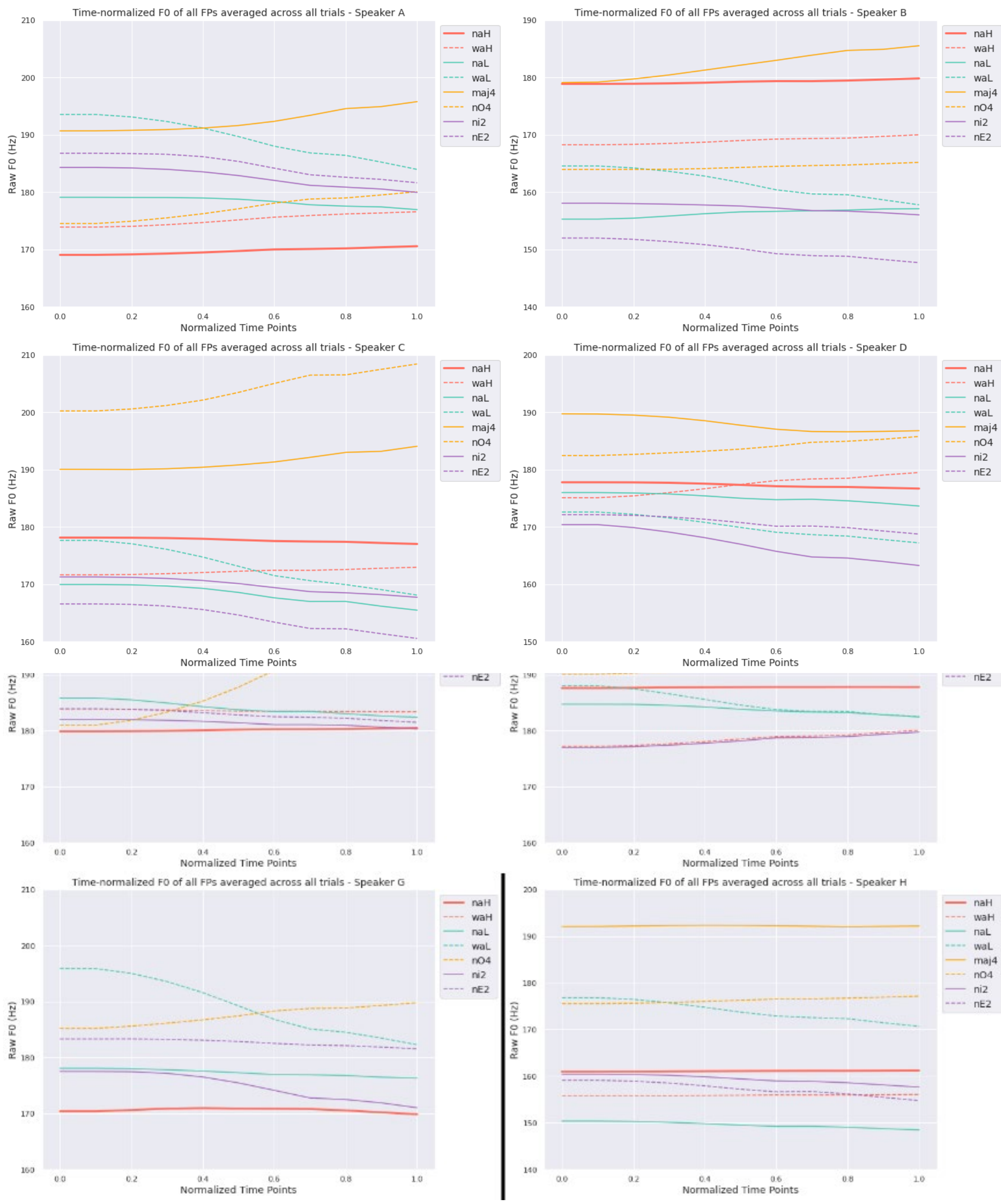


Firstly, there are no prominent inflection points indicative of complex contour tones whether in tonal or toneless FPs. All trajectories are more or less linear, with the only difference being direction and steepness. Tone 4 and H% FPs rise steadily in pitch, with /nɤʔ^4/ standing out as a steeper, more salient climb than other FPs of the same group. Meanwhile, Tone 2 and L% all fall steadily in pitch in a similar manner, both producing moderate movements of minimal deviation in pitch.

However, in terms of location within the pitch range, Tone 4 FPs stand out with their F0 range resting between the 0.2 and 1.3 marks. Their higher range markedly contrasts with Tone 2 and all toneless FPs in the lower range, mostly below the 0.0 mark. In other words, from cross-speaker information, Tone 4 FPs exhibit consistent distinction in pitch range, while other FPs cluster together below said area. Viewing the contours by each individual speaker reveals another distinctive property: Tone 4 FPs consistently reside in a higher pitch range than their Tone 2 counterparts, while toneless FPs vary considerably. For example, the H% FPs of Speakers A, E, and G appear in the lowest range out of all the FPs, despite the alleged 'High' intonational boundary tone target specified in the intonational level. Similarly, L% FPs also have the potential to appear in the middle of the entire pitch range, such can be seen in Speakers A and H. The pitch tracks for individual FPs of each speaker averaged across all trials are provided in Figure 6.

Figure 6

Time-normalized FPs of Speakers A to H, averaged across all trials



In brief, graphical analyses of FPs averaged across all trials and/or speakers demonstrate that toneless and tonal FPs cannot be differentiated through observing points of inflection within the respective contour in order to determine whether it is transitory. Instead, two striking differences are location in the pitch range, where Tone 4 FPs reside consistently higher than other FPs, and greater variability of toneless FPs than of tonal FPs.

In order to quantitatively verify the predicted difference of average ROC between tonal and toneless FPs, Welch's *t*-tests were used to conduct a comparison of means between two independent samples on the average ROC of 1) tonal and toneless FPs, 2) Tone 4 and H% FPs, and 3) Tone 2 and L% FPs.

Tonal vs toneless FPs

Tonal FPs (n = 210, M = 0.51, SD = 0.42) demonstrated a significantly greater ROC than toneless FPs (n = 191, M = 0.40, SD = 0.43), t(394) = -2.78, p < .05. Figure 8 provides a boxplot comparing tonal and toneless FPs back-to-back, with the full Welch's *t*-test report given in Table 3.

Figure 7

Back-to-back boxplots of ROC for toneless and tonal FPs

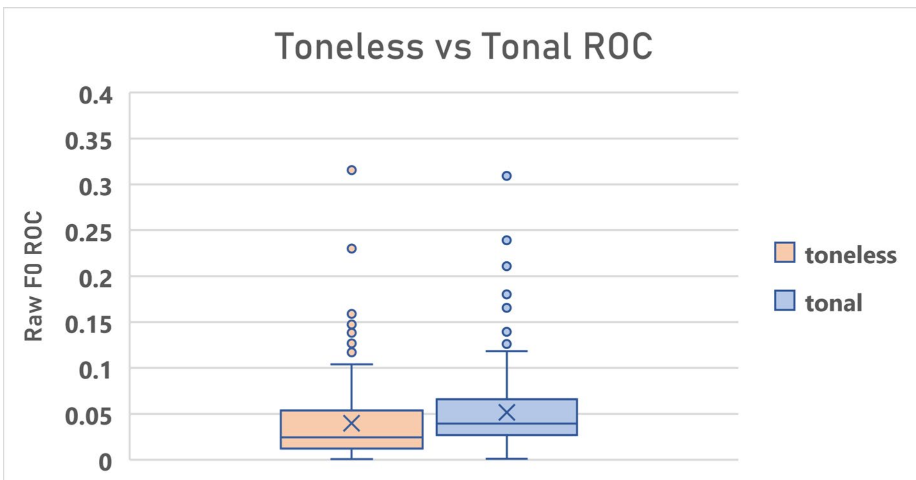


Table 3

Welch's t*-test report for toneless vs tonal FPs*

	Toneless FPs	Tonal FPs
Mean	0.039588153	0.051438
Variance	0.001841271	0.00179
Observations	191	210
df	394	
t Stat	-2.78045	
P(T<=t) two-tail	**0.005689**	
t Critical two-tail	1.966003	

Tone 4 vs H% FPs

Tone 4 FPs (n = 120, M = 0.56, SD = 0.50) demonstrated a significantly greater ROC than H% FPs (n = 100, M = 0.20, SD = 0.17), t(150) = -7.47, p < .05. Figure 9 compares Tone 4 and H% FPs boxplots, with the full Welch's *t*-test report given in Table 4.

Figure 8

Back-to-back boxplots of ROC for H% and Tone 4 FPs

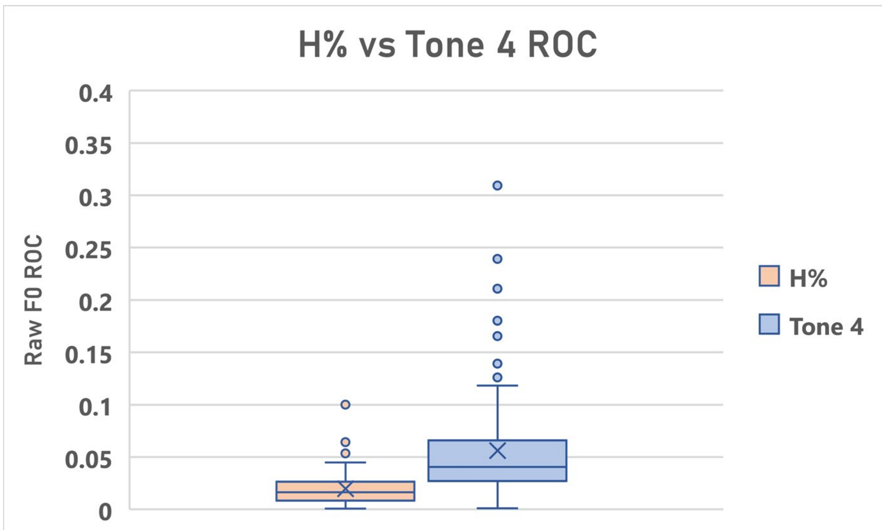


Table 4

Welch's t*-test report for H% vs Tone 4 FPs*

	H% FPs	Tone 4 FPs
Mean	0.019760543	0.056098062
Variance	0.000284993	0.002493866
Observations	100	120
df	150	
t Stat	-7.47487	
P(T<=t) two-tail	**0.000000000597**	
t Critical two-tail	1.975905	

Tone 2 vs L% FPs

Unlike the comparison of Tone 4 FPs with toneless H% FPs, L% FPs (n = 91, M = 0.61, SD = 0.52) demonstrated a significantly greater ROC than Tone 2 FPs (n = 90, M = 0.45, SD = 0.28), t(140) = 2.62, p < .05. Figure 10 compares Tone 2 and L% FPs boxplots, with the full Welch's *t*-test report given in Table 5.

Figure 9

Back-to-back boxplots of ROC for L% and Tone 2 FPs

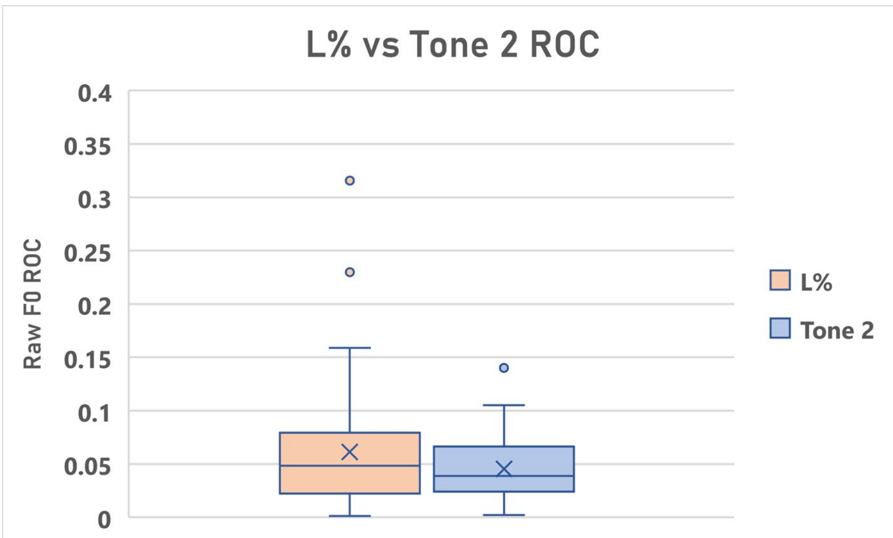


Table 5

Welch's t-*test report for L% vs Tone 2 FPs*

	L% FPs	Tone 2 FPs
Mean	0.061377	0.045224
Variance	0.002657	0.0008
Observations	91	90
df	140	
t Stat	2.617309	
P(T<=t) two-tail	**0.009837**	
t Critical two-tail	1.977054	

In sum, the average ROC values are significantly different for each FP category pair. The ROC values for all tonal FPs are greater than those for all toneless FPs, in line with the narrower comparison between Tone 4 and H% FPs, where Tone 4 FPs alone also exhibit greater rapidity. Surprisingly, the distinction becomes reversed between Tone 2 and L% FPs, where L% FPs show significantly greater ROC values.

Discussion

Recalling the current study's research question, the speech data and the ensuing analyses serve to provide insights into whether tonal and toneless FPs in Thai exhibit similar F0 behavior when occurring inside the intonational phrase. Based on the phonetic characteristics of tone and tonal articulation, lexical tones are expected to exhibit greater movement than the transitory interpolation seen in what is analyzed as toneless syllables (as measured through the average rapidity or ROC).

In terms of cues to the distinction between tonal and toneless FPs, the pitch rate of change seems secondary to pitch range, especially in the amount of relative variability of toneless FPs when compared to the more "fixed-height" tonal FPs. Regardless, quantitatively speaking, ROC does function as a significant cue in differentiating between the two categories, as the ROC values of tonal FPs are significantly greater than those of their toneless counterparts, particularly Tone 4 FPs compared to H% FPs. These findings support the current study's hypothesis that tonal and toneless FPs do not exhibit similar F0 behavior when residing within the intonational phrase.

Prior to concluding that intonational-phrase-internal tonal and toneless FPs are distinct in terms of variability in pitch range and average rate of movement, and that these acoustic disparities reflect their different phonological nature, there is one crucial matter which needs to be addressed. As expected, Tone 4 FPs have a significantly greater ROC than H% FPs. However, Tone 2 FPs exhibit significantly lower ROC values than L% FPs, a complete turnaround from the predictions and the general trend as observed in the speech data. To account for this anomaly, we must return to the notion of tonal target specifications.

Several autosegmental interpretations of Thai tones (e.g. Leben, 1973; Gandour, 1974; Morén & Zsiga, 2006) agree that Tone 2 consists of one autosegment, L, while Tone 4 consists of two autosegments, H and L, regardless of their corresponding TBU. Assuming that H% and L% are both composed of one autosegment, a representational distinction can be observed between Tone 4 vs H% and Tone 2 vs L%. Whereas Tone 4 consists of two tonal targets, Tone 2, H%, and L% all have one, single tonal target. Thus, the FPs specified with the more autosegmentally complex Tone 4 and those specified with their less complex counterpart H% will have more drastic differences in their phonetic realization, as is supported by discovering that Tone 4 FPs have significantly greater ROC than H% FPs. However, the autosegmental specification for Tone 2 and L% are the exact same in quantity and quality, diverging only in their level of representation and occurrence in the surface representation. This similarity might have led to their close resemblances at the phonetic level, susceptible to other external factors beyond controlling for the position within the intonational phrase.

One possible way to account for the unequal difference in ROC between the Tone 4/H% pair and the Tone 2/L% pair may relate to perceptual asymmetry. In Jeon and Heinrich (2022)'s study, it was discovered that pitch height is better discriminated at the peaks (ascending towards a local maximum F0) than at the valleys (descending towards a local minimum F0). Attention is allotted more to pitch movements corresponding to H autosegments than their L counterparts, which means that there is higher perceptual sensitivity in the former case. Despite the work not addressing production itself, "[this perceptual asymmetry] may be linked to how the pitch movements are linguistically coded and tendencies in speech production" (Jeon & Heinrich, 2022).

Stepping back from the theoretical assumptions of this study, it is also possible that FPs, or even most grammatical lexical items, are not specified for tone at all in the lexical representation. Regardless of their phonological conditions and contextual environment, what may appear as a lexical specification of tone could also be the impression of tone applied consciously by the speaker after the fact — especially after being exposed to written stimuli. Apart from revising theoretical concerns, future attempts in answering the relevant research question(s) should consider non-interfering elicitation methods which do not rely on orthographic prompts and retain the desired quantity and quality of the required speech data, all the while controlling for various external factors as discussed above.

Conclusion

According to the AM approach, segmental units without tonal specification realize their pitch through linearly interpolating between the neighboring pitch targets. Such can be predicted for FPs which are structurally analyzed to be "toneless". In this study, it is revealed that the average rate of change values for tonal FPs, especially for Tone 4 FPs when compared to H% FPs, are significantly greater than those for toneless FPs, supporting structural analyses which representationally contrast between the two types of FPs in Thai. As for the difference between Tone 2 and L% FPs, distinctions can be cued through variability in pitch range alone instead of together with rapidity. Future research dealing with the acoustic reality of similarly proposed phonological analyses should keep in mind various theoretical and methodological hurdles discussed within this study.

REFERENCES

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(2), 149-168. https://doi.org/10.1002/wcs.1215

Arvaniti, A., & Fletcher, J. (2020). The autosegmental-metrical theory of intonational phonology. In C. Gussenhoven, & A. Chen (Eds.), *The Oxford handbook of language prosody* (pp. 78-95). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198832232.001.0001

Bee, P. (1975). Restricted phonology in certain Thai linker-syllables. In J. G. Harris & J. R. Chamberlain (Eds.), Studies in Tai Linguistics in Honor of William J. Gedney (pp. 17-32). Central Institute of English Language.

Bolinger, D. (1978). Intonation across languages. In J. Greenberg (Ed.), *Universals of human language* (Vol. Phonology, pp. 471-524). Stanford University Press.

Bradley, C. B. (1911). Graphic analysis of the tone-accents of the Siamese language. *Journal of the American Oriental Society, 31*(3), 282-289.

Choi, J. D. (1995). An acoustic-phonetic underspecification account of Marshallese vowel allophony. *Journal of Phonetics, 23*(3), 323-347. https://doi.org/10.1016/S0095-4470(95)80164-2

Chuenkongchoo, T. (1956). The prosodic characteristics of certain particles in spoken Thai [Master's Thesis, University of London, School of Oriental & African Studies].

Cohn, A. C. (1993). Nasalisation in English: Phonology or phonetics. *Phonology, 10*(1), 43-81. https://doi.org/10.1017/S0952675700001731

Cooke, J. R. (1989). Thai sentence particles: Forms, meaning and formal-semantic variations. In *Thai sentence particles and other topics* (pp. 1-90). Pacific Linguistics.

Faitaki, F., & Murphy, V. A. (2020). Oral language elicitation tasks in applied linguistics research. In *The Routledge handbook of research methods in applied linguistics* (pp. 360-369). Oxfordshire, UK: Routledge.

Frazier, K. (2014). Implications of autosegmental analysis in the exploration of prosodic phonology in Mandarin Chinese [Master's thesis, Liberty University]. https://digitalcommons.liberty.edu/honors/483

Gandour, J. (1974). On the representation of tone in Siamese. In J. G. Harris & J. R. Chamberlain (Eds.), Studies in Tai Linguistics in Honor of William J. Gedney (pp. 170-195). Central Institute of English Language.

Gandour, J., Potisuk, S., & Dechongkit, S. (1994). Tonal coarticulation in Thai. *Journal of Phonetics*, *22*, 477-492. https://doi.org/10.1016/S0095-4470(19)30296-7

Gandour, J., Tumtavitikul, A., & Satthamnuwong, N. (1999). Effects of speaking rate on Thai tones. *Phonetica, 56*(3-4), 123-134. https://doi:10.1159/000028447

Greif, M. (2010). Tones and intonation in Prinmi – a first survey. *STUF - Language Typology and Universals, 63*(3), 221'251. https://doi.org/10.1524/stuf.2010.0020

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge University Press.

Ha, K. P. (2012). *Prosody in Vietnamese: Intonational form and function of short utterances in conversation*. University of Cologne

Hardcastle, W. J., & Hewlett, N. (2006). Introduction. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 1-5). Cambridge University Press.

Henderson, E. J. (1949). Prosodies in Siamese: A study in synthesis. *Asia Minor, 1*, 189-215.

Hyman, L. M. (2001). Tone systems. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals* (Vol. 2, pp. 1367-1379). New York, NY: Walter de Gruyter.

Hyman, L. M. (2006). Elicitation as experimental phonology: Thlantlang Lai tonology. *UC Berkeley PhonLab Annual report, 2*. https://doi.org/10.5070/P74nr5n1z0

Jeon, H.-s., & Heinrich, A. (2022). Perceptual asymmetry between pitch peaks and valleys. *Speech Communication, 140*, 109-127. https//doi.org/10.1016/j.specom.2022.04.001

Joshua Project. (2021). *Language: Ruund.* https://joshuaproject.net/languages/rnd

Kallayanamit, S. (2004). Intonation in Standard Thai: Contours, registers and boundary tones (Publication No. 3148629) [Doctoral dissertation, Georgetown University]. ProQuest Dissertations & Theses Database.

Keane, E. (2006). Prominence in Tamil. *Journal of the International Phonetic Association, 36(1)*, 1-20. https://doi.org/10.1017/S0025100306002337

Keane, E. (2012). The intonational phonology of Tamil. In S.-a. Jun (Ed.), *Prosodic typology II: The phonology of intonation and phrasing* (pp. 118-153). doi:10.1093/acprof:oso/9780199567300.003.0005

Keating, P. A. (1988). Underspecification in phonetics. *Phonology, 5*(2), 275-292. https://doi.org/10.1017/S095267570000230X

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.

Law, S.-p. (1990). The syntax and phonology of Cantonese sentence-final particles (Publication No. 9022980) [Doctoral dissertation, Boston University]. ProQuest Dissertations & Theses Database.

Leben, W. (1973). Suprasegmental phonology [Doctoral Dissertation, Massachusetts Institute of Technology].

Lionnet, F., & Hyman L. M. (2018). Current issues in African phonology. In T. Güldemann (Ed.), *The languages and linguistics of Africa* (pp. 602-708). https://doi.org/10.1515/9783110421668-006

Luksaneeyanawin, S. (1983). Intonation in Thai [Doctoral dissertation, The University of Edinburgh]. Edinburgh Research Archive. http://hdl.handle.net/1842/6722

Luksaneeyanawin, S. (1998). Intonation in Thai. In D. Hirst, & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 376-394). Cambridge University Press.

McPherson, L. E. (2011). Tonal underspecification and interpolation in Tommo So [Master's thesis, University of California, Los Angeles]. https://linguistics.ucla.edu/general/matheses/McPherson_UCLA_MA_2011.pdf

McPherson, L. E. (2013). *A Grammar of Tommo So*. De Gruyter Mouton.

Morén, B., & Zsiga, E. (2006). The lexical and post-lexical phonology of Thai tones. *Natural Language & Linguistic Theory, 24*(1), 113-178. https://doi.org/10.1007/s11049-004-5454-y

Moussa, H. A. (2019). Jeddah Arabic intonation: An autosegmental-metrical approach [Master's Thesis, Newcastle University]. http://theses.ncl.ac.uk/jspui/handle/10443/5386

Noss, R. (1964). *Thai reference grammar*. Foreign Service Institute.

Odden, D. (1982). Tonal phenomena in Kishambaa. *Studies in African Linguistics, 13*(2), 177-208.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation [Doctoral dissertation, Massachusetts Institute of Technology].

Pittayaporn, P. (2014). Prosody of final particles in Thai: Interactions between lexical tones and boundary tones [Manuscript in preparation].

Ross, E. D., Edmondson, J. A., & Seibert, G. B. (1986). The effect of affect on various acoustic measures of prosody in tone and non-tone languages: A comparison based on computer analysis of voice. *Journal of Phonetics, 14*, 283-302.

Welmers, W. E. (1959). Tonemics, morphotonemics, and tonal morphemes. *General Linguistics, 4*(1), 1-9.

Wong, W. P., Chan, M., & Beckman, M. E., (2005). An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In S.-a. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 271-299). https://doi.org/10.1093/acprof:oso/9780199249633.003.0010

Xu, Y. (2005-2021). ProsodyPro. http://www.homepages.ucl.ac.uk/~uclyyix/ProsodyPro/

Yip, M. (1980). The tonal phonology of Chinese [Doctoral dissertation, Massachusetts Institute of Technology]. DSpace@MIT. http://dspace.mit.edu/handle/1721.1/15971

Yip, M. (2002). *Tone*. Cambridge University Press.

Zhang, J. (2018). A comparison of tone normalization methods for language variation research. *Proceedings for the 32nd Pacific Asia Conference on Language, Information and Computation*, (pp. 823-831). Hong Kong. https://aclanthology.org/Y18-1095.pdf

Acknowledgments

I would like to extend my gratitude to Dr. Pittayawat Pittayaporn for his extensive mentorship, thoughtful advice, and incredible support, both academic and personal, throughout the research project. Without his insightful supervision, this project would never have been completed. I also thank the anonymous reviewers for their detailed and thoughtful comments and points of discussion. Furthermore, I am indebted to the faculty members and peers at the Department of Linguistics, Chulalongkorn University, for their generous words along the way. Finally, I give my thanks to all the participants who took part in the experiment.

Appendix A

The demographic and linguistic background information for each participant can be accessed via the following link:

https://docs.google.com/spreadsheets/d/1eX6dgJsNVT4vLkPGpx5g15EK85T6Dc4h36nqFhGG0jY/edit?usp=sharing

Appendix B

The script for each dialogue variant can be accessed via the following link:

https://drive.google.com/file/d/1reYU6jiqsZXiUJ7WvR5c6oXn8BLzNPkQ/view?usp=sharing

Appendix C

The pitch tracks of individual FPs of each speaker averaged across all trials can be accessed via the following link:

https://drive.google.com/drive/folders/1GDfpr7y-wWtxlN1o-T8q946LQy_TWTFA?usp=sharing

Appendix D

The Google Colaboratory script for pitch track visualization can be accessed via the following link:

https://colab.research.google.com/drive/1dBgwkFNsSNUXFPWdc1LOMsgPnWcmpb3Z?usp=sharing