Journal of the Southeast Asian Linguistics Society
JSEALS 17.1 (2024): 63-80
Research Article
ISSN: 1836-6821, DOI: http://hdl.handle.net/52523
University of Hawai'i Press

IU MIEN TONE CHANGE IN REAL TIME: A RESTUDY OF L-THONGKUM (1988)

Cathryn Yang
Payap University
cathryn_yang@sil.org

Jennifer Engelkemier
Payap University
jennifer_e@payap.ac.th

Tatsuro Daniel Arisawa
Iu Mien Mother-Language Preservation and Development Center
tdanielari@gmail.com

Abstract
This study investigates tone changes in Iu Mien, a Hmong-Mien language with two complex contour tones: high rising-falling (T3) and low rising-falling (T4). L-Thongkum (1988) observed younger speakers producing rising variants of T3 and T4, which she attributed to Thai contact. This study replicates L-Thongkum's study 34 years later to observe tone changes in real time and examine the potential role of peak delay, a phonetic bias resulting in a later f0 peak. 40 speakers (ages 12-84) produced T3 and T4 monosyllables in isolation and in a carrier phrase. f0 was modeled using Generalized Additive Mixed Modeling (GAMM). Results confirm incremental change toward rising variants of T3 and T4. Older speakers showed peak delay (with more of a rising contour) under conditions of shorter duration and preceding low tone. Critically, younger speakers produced peak-delayed, rising variants in all conditions, indicating generalization of the rising contour. While language contact may have initiated the changes, a phonetic bias toward peak delay generated the seeds of potential change. The change from rising-falling > rising aligns with apparent-time studies in other Asian languages, suggesting this may be a crosslinguistically unidirectional pathway.

Keywords: tone change, real time, peak delay
ISO 639-3 codes: ium, tha

1 Introduction

Research on diachronic tone change lags behind segmental sound change (Thomas 2019; Campbell 2021). The constraints problem, i.e., delineating limitations on direction of change (Weinreich, Labov & Herzog 1968), is one of the foundational questions in the study of sound change, and yet remains underexplored for tone. Until recently, the lack of acoustic tone analysis from earlier time periods hindered the documentation of tone change over time (Pittayaporn 2018). Researchers were limited to apparent-time studies, where change is inferred based on generational differences observed at one point in time (Labov 1994). Fortunately, the instrumental study of tone began in the last century (Bradley 1911), making it now possible to compare acoustic results for certain languages across time periods and observe change in real time. The field has entered an exciting era of discovery.

The present study examines tone variation and change over the past 34 years in the Iu Mien language (ISO 639-3 code [ium]; Glottocode [iumi1238]) of northern Thailand. Iu Mien is a Hmong-Mien language spoken in southern China and surrounding areas by over 800,000 speakers; there are approximately 45,000 speakers in Thailand (Eberhard, Simons & Fennig 2023). L-Thongkum (1988) conducted one of the first acoustic studies of Iu Mien tone, recording speakers from various age groups in Huai Mae Sai Village, Chiang Rai Province. Based on data from the 1988 study, L-Thongkum (1997) identified generational differences in the pronunciation of Tone 3 (T3; high rising-falling) and Tone 4 (T4; low rising-falling). Specifically, L-Thongkum (1997) observed that younger speakers' pronunciations of T3 and T4 had become phonetically similar to the High and Rising tones of Standard Thai, which she interpreted as the result of ongoing change due to contact with the national language.

We replicate L-Thongkum's (1988) study with an apparent-time study conducted in 2022, 34 years later, and compare our results to hers. The present study is one of the first trend studies on tone change, enabling observations in real time. Our restudy of the Iu Mien community of Huai Mae Sai, Thailand, has the potential to confirm L-Thongkum's (1997) interpretation of generational differences as reflecting ongoing tone change. Clearly establishing the direction of tone change contributes towards solving the constraints problem by providing empirical evidence of the pathways along which tones may shift over time. In addition to documenting change in real time, we also aim to investigate the role of the phonetic bias known as "peak delay" (Xu 2001) in facilitating the observed tonal changes. Peak delay refers to a rightward shift in the alignment of an f0 peak, which can cause rising-falling tones to take on more of a rising shape over time. Examining whether peak-delayed variants expand their distribution across generations can shed light on how phonetic bias may result in diachronic tone change.

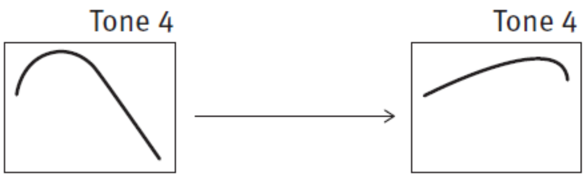
Our approach aligns with sound change models that aim to connect synchronic phonetic and systemic biases with diachronic change (Blevins 2004; Garrett & Johnson 2013; Ohala 2003; Pittayaporn 2018). Crosslinguistically common sound changes are often unidirectional: they happen in one direction but not in the opposite direction (Bybee 2012). For example, velars often become palatals before fronted vowels, but palatals perhaps never become velars in such conditions (Harrington et al. 2019). Such directionality is thought to reflect the fact that phonetic variants—the seeds of potential sound change—are not generated randomly but rather through production and perception biases (Garrett & Johnson 2013; Pittayaporn 2018).

Almost all models of sound change agree that synchronic variation is a necessary (but not sufficient) condition for diachronic sound change (Bybee 2012; Chitoran 2012; Bermúdez-Otero 2020; Yu 2021). Sound change occurs when a phonetic variant becomes more frequent, expanding its distribution and shifting the center of the range of variation (Bybee 1999). Bermúdez-Otero (2020) proposes that sound change may happen at the community level if members perceive an inverse correlation between the use of a variant and age (an "age vector") and then align their use according to that age vector ("momentum-sensitive learning"). Momentum-sensitive learning may account for why phonetic biases continually generate contextual variants, but only some of those variants undergo community-level propagation (Bermúdez-Otero 2020).

An important parameter of tone variation is tonal alignment, the timing relationship between a tone and its tone-bearing-unit (TBU), the portion of the segmental string that the tone is associated to (Karlin 2018; Barnes et al. 2012). A tone's f0 movement unfolds in time simultaneously with the segmental string, and tonal alignment conveys crucial cues for tone perception (Burroni 2023; Zsiga & Nitisaroj 2007). One articulatory constraint on tonal alignment is the maximum speed of f0 change, which may be approached even at normal speech rates (Xu & Sun 2002). This physiological limitation can result in peak delay, which has been defined as when the f0 peak occurs in the syllable following the one it is lexically associated with (Xu 2001). In this paper, we broaden the definition of peak delay to include peaks shifting to a later alignment even within the lexically-associated syllable, not only onto the following syllable.

Based on the bias toward synchronic peak delay, Pittayaporn (2018: 262) predicts that, diachronically, f0 peaks are more likely to slide rightwards than leftwards, a directionality he terms "peak sliding." Pittayaporn (2018) illustrates peak sliding through a case study of change in real time in Bangkok Thai Tone 4 "High", which morphed from having an early-aligned peak to a late-aligned one during the first half of the 20^{th} century (see Figure 1). Note that when a peak slides to the right, overall contour shape changes as well, as in Thai Tone 4 moving from a rising-falling contour to more of a rising contour.

Figure 1: *Peak sliding in Tone 4 "High" of Bangkok Thai in the first half of the 20^{th} century, adapted with permission from Pittayaporn (2018) Figure 5.*



Crosslinguistically, rising-falling tones may be susceptible to peak sliding. Yang & Xu's (2019) survey of tone change in Asian languages found several instances of contour change in rising-falling tones that appeared to reflect peak sliding—i.e., the peak of a rising-falling tone shifting to a later alignment. In contrast, instances of a peak shifting to an earlier alignment were not reported. This finding suggests that rightward sliding (as opposed to leftward sliding) is a unidirectional trend, at least when affecting rising-falling tones.

The phonetic bias for peak delay may be a contributing factor in the specific changes observed in the Iu Mien rising-falling tones T3 and T4. L-Thongkum (1997) attributed the T3 and T4 changes to language contact, without considering the role of phonetic bias. This study expands on the previous study by also examining language-internal factors, specifically peak delay, since rising-falling tones appear prone to peak sliding.

Peak delay is likely to occur at faster speech rates, when the rate of syllable production exceeds the physiological limit for rapid pitch changes (Xu 2001). We therefore predict that peak delay, broadly defined, will probably occur in target syllables produced at a faster speech rate. Peak delay also tends to occur when the preceding tone's f0 offset is low but the target tone's peak is high, since a larger f0 transition requires a longer time (Xu & Sun 2002). We therefore predict that a preceding low tone will also favor peak delay. We especially want to know if peak-delayed variants expanded in their distribution across generations, i.e., whether younger speakers produce rightward shifted variants not only in the original conditioning environment but also in other environments.

The present study addresses two research questions concerning variation and change in T3 and T4:
1) Comparing the 1988 and 2022 acoustic results, is there evidence of a monotonic (unidirectional) trajectory of change in real time?
2) Does peak delay occur under conditions of preceding low tone and shorter duration, and does the peak-delayed variant expand its distribution across generations?

2 Variation and change in Iu Mien tones

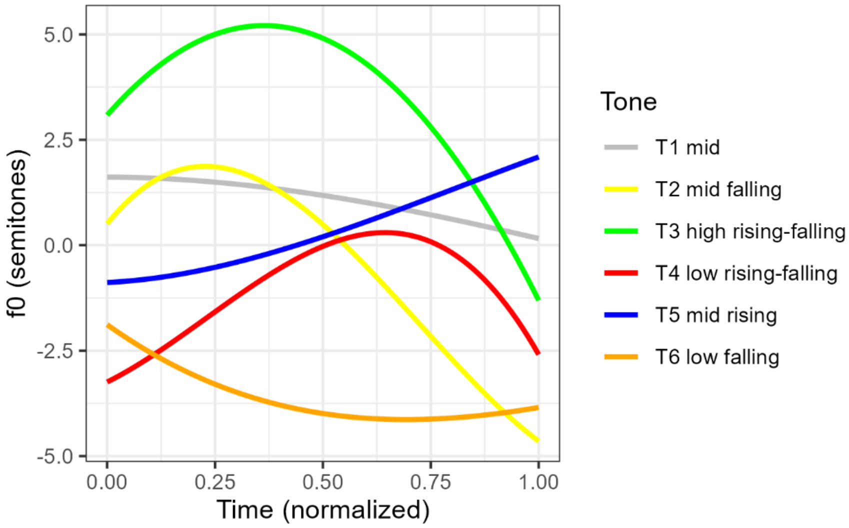
Iu Mien belongs to the Mienic branch of the Hmong-Mien family (Ratliff 2010). Most words are monosyllabic, and syllable structure is (C)V(V)(C) (Prasertsud 2017). Purnell (2012:xxv) describes two tone alternation patterns in disyllabic words, but since the present study's focus is monosyllabic words, we do not discuss these alternations further. L-Thongkum (1997) analyzes Iu Mien as having six tones, described in Table 1:

Table 1: *Description of the Iu Mien tone system according to L-Thongkum (1997:156)*

Tone label	Orthographic marker	Historic tone category	Chao pitch number	Prose description
T1	unmarked	A1	33	mid level
T2	-h	A2	31	mid falling
T3	-v	B1	453	high rising-falling
T4	-z	B2	232	low rising-falling
T5	-x	C1	34	mid rising
T6	-c	C2	21	low falling
T3 checked	-v	D1	45	high rising
T6 checked	-c	D2	21	low falling

L-Thongkum (1988) recorded 24 Iu Mien speakers (age range 17-56 years; 12 female) in Huai Mae Sai Village. Her report provided f0 measurements at the 0%, 25%, 50%, 75%, and 100% timepoint for one token per tone per speaker. Figure 2 presents acoustic analysis results from the 1988 dataset based on six speakers born in the 1940's. F0 is converted to semitones relative to each speaker's f0 mean (the zero level in Figure 2; Zhang 2018) and fitted to a cubic polynomial curve using the geom_smooth function in R (Wickham 2016; R Core Team 2023). The curves in Figure 2 match the prose descriptions in Table 1.

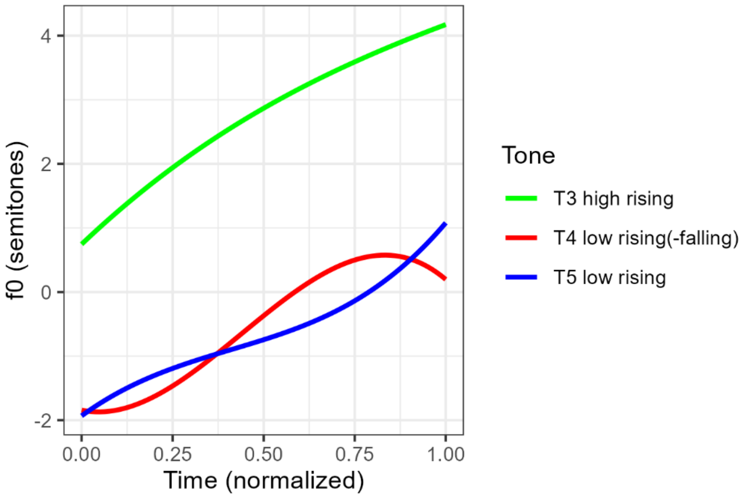
Figure 2: *Iu Mien tone system, redrawn from L-Thongkum (1988)'s dataset, using a subset of six speakers (three female) born in the 1940's, fit to a polynomial curve with the geom_smooth function in ggplot2 (Wickham 2016)*



L-Thongkum (1997) described generational differences in the pronunciation of T3 and T4, claiming that younger speakers pronounced T3 as a high rising tone and T4 as a low rising tone. Figure 3 shows the acoustic analysis results from a female speaker born in the 1970's from L-Thongkum (1988)'s study. For this speaker, the contour of T3 is simply rising, and the final fall of T4 is late in the syllable and only slight. Also, T5 appears to have lowered from mid rising to low rising. Changes to T4 and T5 result in increased phonetic similarity between these two contrastive tonemes. Based on these findings, L-Thongkum (1997) predicted that T4 and T5 would merge completely, resulting in a ~~five~~ five-tone system. She claimed that the observed generational differences in T3 and T4 were due to contact with Standard Thai.

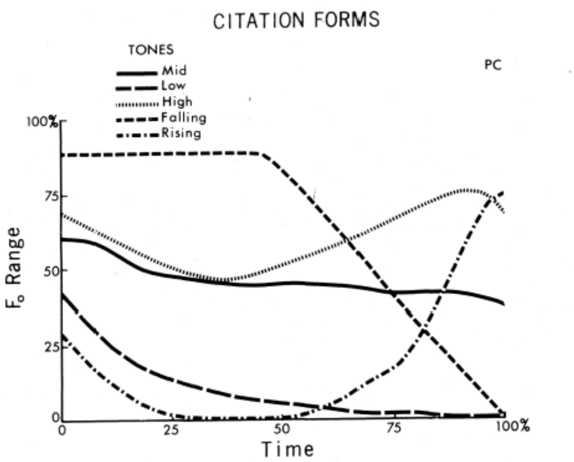
Increased proficiency in Standard Thai among younger Iu Mien speakers is the basis for L-Thongkum (1997)'s claim. L-Thongkum (1997) described the rapidly changing sociolinguistic context of the Iu Mien community in Huai Mae Sai Village in the late 1980's and early 1990's. A primary school built in the village began offering basic education in 1978, with Standard Thai as the medium of instruction. Increased access to education, as well as a growing tourism industry, led to a dramatic increase in proficiency in Standard Thai among young people by the early 1990's (p. 154). However, intra-group interaction was still reserved for Iu Mien, with Northern Thai spoken with local non-Iu Mien neighbors and Standard Thai spoken in formal situations (p. 155).

Figure 3: *Iu Mien Tones 3, 4, and 5, redrawn from L-Thongkum (1988)'s dataset, using a subset of one female speaker born in the 1970's, fit to a polynomial curve with the geom_smooth function in ggplot2 (Wickham 2016)*



To contextualize L-Thongkum's (1997) claim of contact-induced change, we now illustrate Standard Thai tone pronunciation in the 1970's, when the first school cohort in Huai Mae Sai was educated in Standard Thai. Figure 4 shows tones produced by a Bangkok Thai speaker recorded in 1973-1974 (Abramson 1979). Standard Thai has five tone contrasts, commonly labeled as Mid, Low, Falling, High, and Rising (Abramson 1979). L-Thongkum (1997) describes Standard Thai High Tone as a high rising tone (Chao pitch numbers [45]) and Rising Tone as a low rising tone (Chao numbers [15]). In Figure 4, High and Rising appear with a rising contour, albeit with an initial f0 dip before the rise begins. The high rising contour that younger Iu Mien speakers produced for Tone 3 in L-Thongkum (1997) therefore matched Thai High Tone's rising contour. Similarly, the low rising variant of Tone 4 produced by younger Iu Mien was like the Rising tone in Thai. The changes of Iu Mien T3 and T4 from rising-falling to rising thus result in a tone system that is more phonetically similar to Standard Thai. Also, the merger of T4 with T5 would result in a five-tone system, matching the inventory size of Standard Thai. Consequently, L-Thongkum (1997) described these tonal shifts as being contact-induced.

Figure 4: *Average f0 contours on isolated Thai monosyllables for one speaker. Reproduced from Abramson (1979) Figure 1.*



3 Methodology

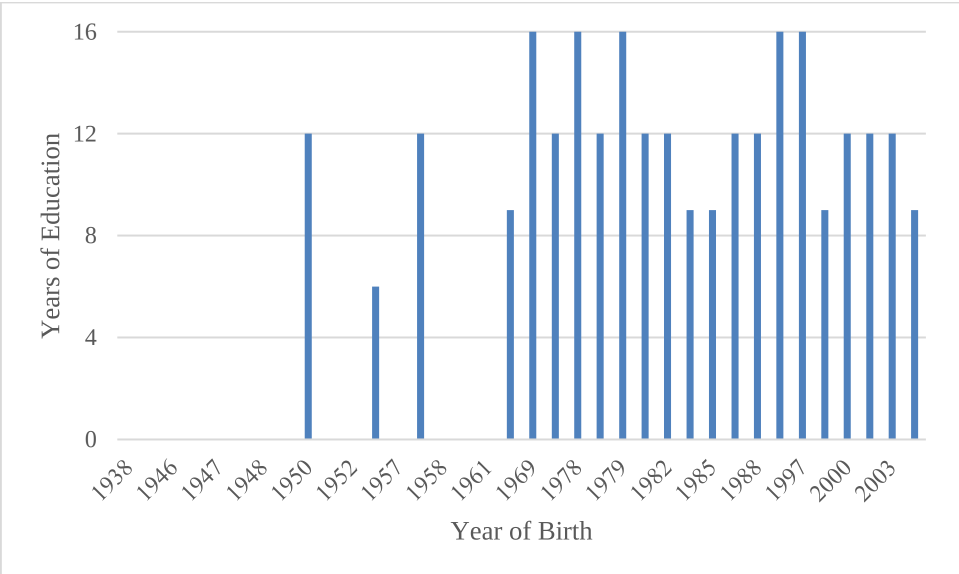
3.1 Data collection

The current study employed a methodology designed to elicit production data and facilitate comparison with previous research. A picture-naming task rather than a reading task was chosen to elicit production, to minimize potential influence from reading a Thai-based script. The list of items to be named (see Appendix in supplementary materials) consisted of 50 monosyllabic words taken from everyday life. The 50-item list included all 24 words listed in L-Thongkum (1997), in order to increase comparability between the present study and L-Thongkum (1997). The 7-10 words per tone in the list were roughly balanced for syllable-initial consonant type (voiceless obstruent, voiced obstruent, voiced sonorant). For each word, a line drawing was selected as prompt in the naming task.

Forty native speakers of Iu Mien were recruited for recording (20 female, ages 12-74, median 44.5; 20 male, ages 14-84, median 48). Speakers were recorded in two locations: Huai Mae Sai Village, the field site of L-Thongkum's (1988, 1997) earlier studies, and Huai Kaeo Village, located within 20 kilometers of Huai Mae Sai. Twenty-two participants reported Huai Mae Saai Village as their hometown. Other participants reported locations such as Huai Kaeo and Huai Chomphu, which are considered to belong to the same dialect as Huai Mae Saai. All speakers reported fluency in Northern Thai, Standard Thai, and Iu Mien.

Compared to L-Thongkum's (1988, 1997) earlier studies, one notable sociolinguistic difference is the increased use of Thai (Standard and Northern) among the 2022 sample. All participants aged 55 and older reported Iu Mien as the language they speak most fluently, but almost all participants under the age of 55 reported Thai as their favored language, with Iu Mien listed as second. This generational difference reflects increased access to Thai through formal education. Figure 5 illustrates the strong relationship between year of birth and education level in the 2022 sample, with participants born after 1970 attaining much higher levels of education in the Thai language compared to previous generations.

Figure 5: *Education levels by year of birth in the 2022 sample*



The data collection procedure was conducted in Standard Thai and Iu Mien. After orientation to the research and an interview about the speaker's demographic information, the speaker was asked to look at each prompt (a line drawing) and give the name of the item in Iu Mien. Speakers were instructed to repeat the word four times and once in a carrier sentence.

The carrier sentence in (1) embedded the target word between two low tones. This phrase-medial context is favorable to peak delay for two reasons. First, a preceding low tone requires a longer transition to a high f0 peak (Xu & Sun 2002). Second, in connected speech the syllable duration is typically shorter than in isolation (Xu 2006). Comparison of isolation and phrase-medial forms enables us to address RQ2, which examines whether the distribution of peak-delayed variants increases across generations.

(1)	naːi˧˥	buə˧	heu˩		o˩
	Naaiv	mbuo	heuc	_____	oc
	this	1P.inc	call		polite.particle
	"This we call a ____ ."				

Recordings were made with a Zoom H4nPro recorder and Shure SM35 headset condenser microphone positioned approximately one inch from the participants' mouths (sampling rate of 44.1 kHz). Each participant received a small gift of money and/or boxed soymilk in thanks for their time (20-30 minutes per participant).

3.2 Data processing and analysis

The rhyme (nucleus and coda) of each token was manually segmented and labeled in Praat (Boersma & Weenick 2023). F0 values were extracted every 10 msec (Stanford 2016). For cross-speaker comparison, f0 was converted to semitones in reference to speaker-specific mean f0 (Zhang 2018). Tokens that had measurement errors or were shorter than 50 milliseconds were excluded, leaving a total of 1777 T3 tokens (1443 citation form, 334 phrase-medial form) and 1416 T4 tokens (1144 citation form, 272 phrase-medial form). Extracted f0 data and the code used to analyze and visualize the data are available in the supplementary materials repository at https://doi.org/10.5281/zenodo.10597579.

We explored the effect of language-internal and language-external factors on f0 for T3 and T4 using general additive mixed modeling (GAMM). GAMMs are like linear mixed effects regression modeling in that both can include fixed and random effects, hence "mixed" (Sóskuthy 2017). One difference is that GAMMs model non-linear relations between predictors and outcome by including "smooth" terms (Sóskuthy 2021; Wieling 2018). This is useful when modeling dynamic tonal contours, because f0 varies non-linearly with time across the syllable (Chuang, Fon & Baayen 2021). R packages mgcv version 1.9-0 (Wood 2017) and itsadug version 2.4.1 (Rij et al. 2022) were utilized for model fitting, data visualizations and model comparison, following recommendations from Sóskuthy (2017).

We constructed one model for T3 and another for T4 and T5 together, since we are interested in potential merger between T4 and T5 (though discussion of merger falls outside the scope of this paper). After inspection of the distribution of the residuals revealed two heavy tails (i.e., a t-distribution rather than a normal distribution), we modeled the data with a scaled t-distribution (family = "scat"), following Chuang et al. (2021). In both models, f0 is the outcome variable. Sex (*male, female*) and Context (*phrase-medial, citation*) were included as parametric terms to capture variation in overall f0 height.

The data collected in the present study is not sufficient to tease apart the impact of contact versus phonetic bias. Variation in Thai proficiency across speakers could potentially shed light on the role of language contact in the changes, but in the 2022 sample, education level is colinear with age, such that speakers born after 1970 attained much higher levels of education in the Thai language compared to previous generations (see Figure 5). As a result, using education as a proxy for Thai contact is not feasible, as its effect cannot be disentangled from age. The focus of this study was therefore on documenting any incremental changes across generations and examining potential phonetic biases like peak delay that may have influenced the direction of the change. Disentangling language-internal and language-external forces will require alternative approaches in future research.

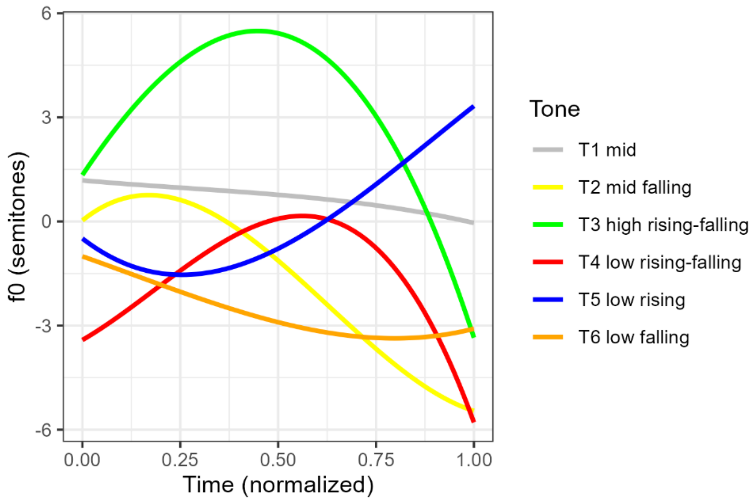
To model the dynamic shape of the contour, a smooth term of Timepoint (centered) was incorporated as a continuous variable. Additional smooth terms for Decade of birth and Duration were included. To investigate the effect of Decade of birth on contour shape, a tensor product interaction between Timepoint and Decade of birth was included. A tensor product interaction allows the shape of the non-linear smooth (here representing the tonal contour over the duration of the rhyme) to vary by another factor (here, Decade of Birth). This enables examining whether the contour shape changes systematically across generations. Similarly, a tensor product interaction between Timepoint and Duration was included to examine the effect of duration on contour shape. The model also contained a tensor product interaction between Duration and Decade of birth, to examine how the effect of Duration on f0 was modulated by Decade of birth. All smooth terms and tensor product interactions were allowed to vary by Sex and by Context. To account for individual speaker and word variability, random smooths for Speaker and Word were included.

The T4-T5 model was identical to the T3 model in structure except all smooth terms were additionally allowed to vary by Tone (two levels: Tone 4 or Tone 5). Model comparison was accomplished by creating a model that excluded the parametric and smooth terms corresponding to a predictor, as well as interactions involving that predictor, and then comparing that model to the full model using ANOVA (Sóskuthy et al. 2018).

4 Results

Figure 6 presents a tone plot of the Iu Mien tone system as produced by speakers recorded in 2022 who were born in the 1940's, representing a conservative pronunciation of the tones. The basic shapes of all the tones are similar to the tone plot seen in Figure 2, which represents the pronunciation of speakers born in the 1940's recorded in 1988. The minimal differences between 1988 dataset and the current dataset suggests that speakers born in the 1940's generally have remained stable in their pronunciation of tone over the past 34 years.

Figure 6: *Iu Mien tone system based on data from the current study, using a subset of seven speakers (two female) born in the 1940's, fit to a polynomial curve with the geom_smooth function in ggplot2 (Wickham 2016)*



4.1 RQ1: Trajectory of change in real time

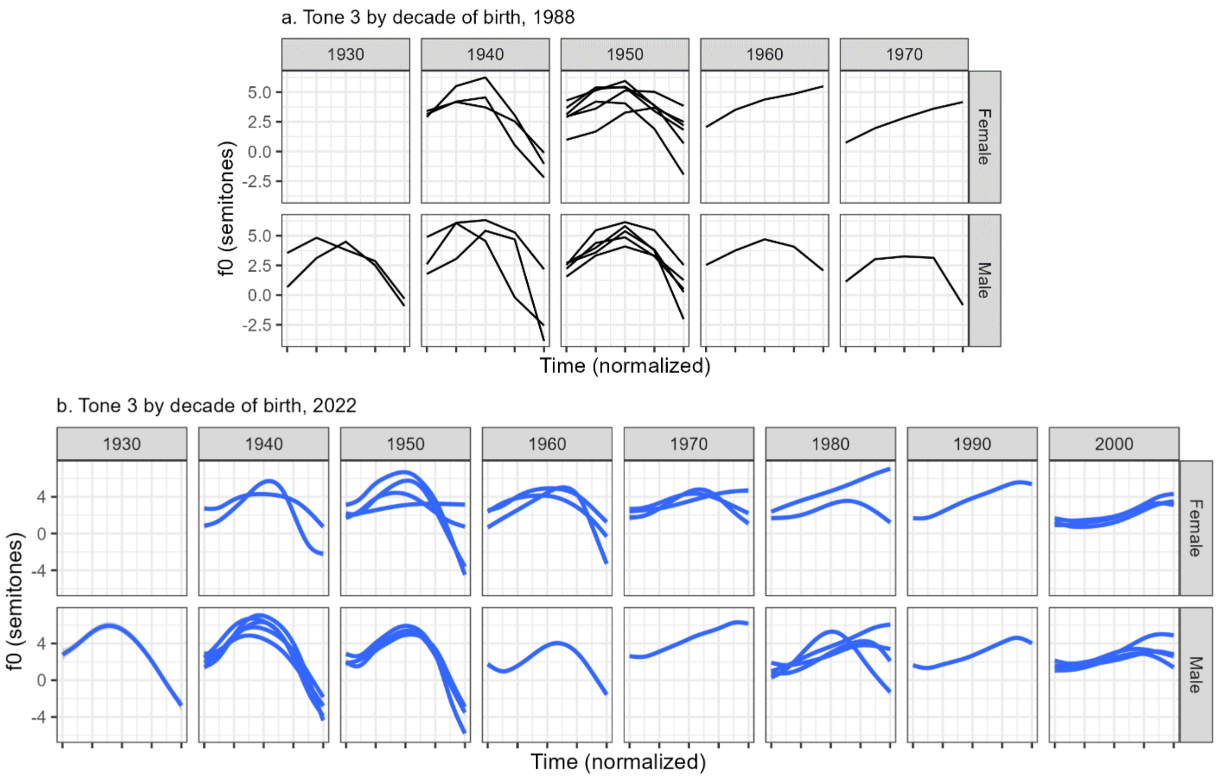
In this section, we address RQ1: is there evidence of a consistent trajectory of change in real time? We compare L-Thongkum's (1988) results with our 2022 results to examine whether the change rising-falling > rising shows gradual incrementation across generations. This section allows visual inspection for trends across generations, and Section 4.2 presents the results of the GAMM modeling.

4.1.1 Tone 3

We first examine generational variation in Tone 3, the high rising-falling tone, as shown in Figure 7. Each individual trajectory in Figure 7 represents the pronunciation of one speaker: the 1988 dataset provided data of only one citation token per speaker, but the 2022 dataset provides an average of 36 citation tokens per speaker, which are then fit to one curve per speaker using the geom_smooth function in R (Wickham 2016; R Core Team 2023).

A comparison of Figure 7a (from 1988) and 7b (from 2022) supports L-Thongkum's claim regarding the direction of change. Figure 7a shows that speakers born before 1960 consistently produced rising-falling, but two female speakers born in the 1960's and 1970's produced rising. In Figure 7b (the 2022 dataset), the majority of speakers born after 1970 produce rising. Though there are a few speakers born in the 1970's and 1980's that maintain rising-falling, almost all speakers born after 1990 produce a rising contour.

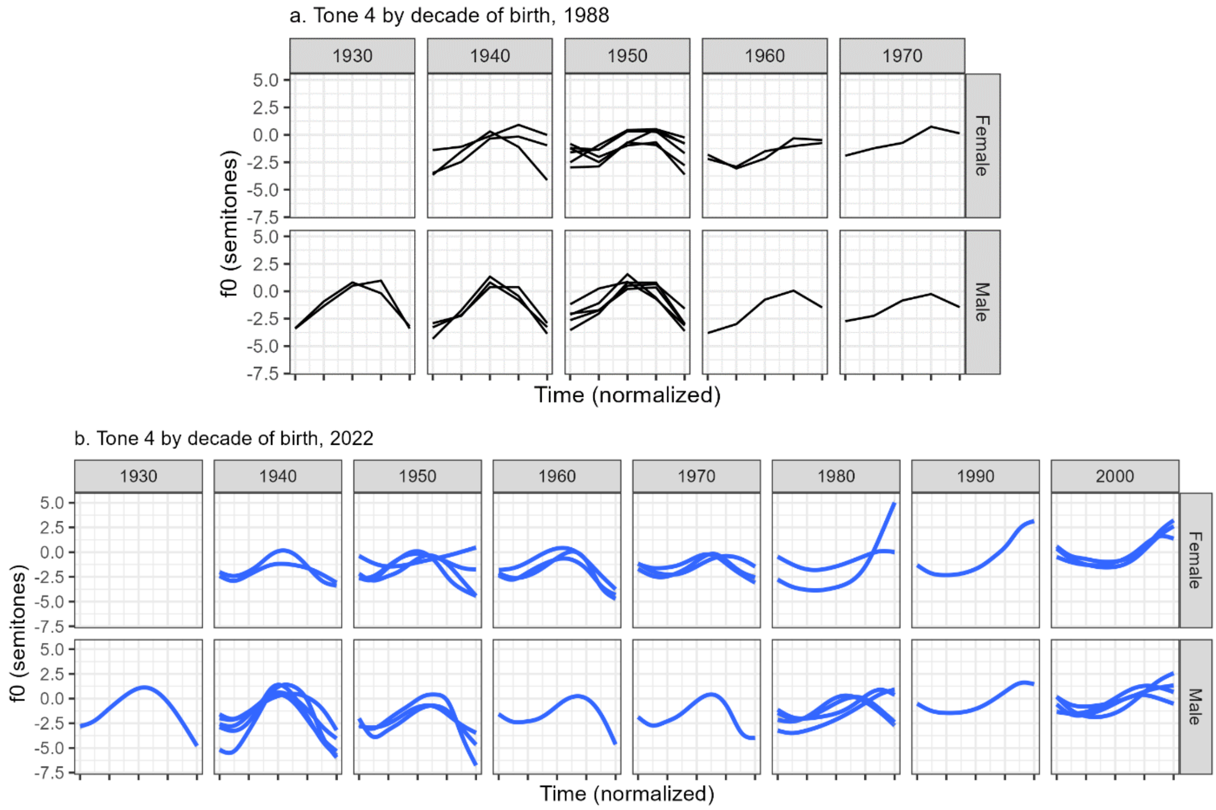
Figure 7a) *Tone 3 f0 trajectory for each speaker from L-Thongkum (1988)'s dataset, which provided one token per speaker, drawn with geom_line function in ggplot2 (Wickham 2016);* ***7b)*** *Tone 3 f0 trajectory for each speaker from the current dataset, with multiple tokens fit to one curve per speaker with the geom_smooth function "gam" method in ggplot2 (Wickham 2016)*



4.1.2 Tone 4

A trend similar to the results described for Tone 3 is also seen for Tone 4. Figure 8 presents speakers' Tone 4 trajectories for the 1988 dataset (one token per speaker) vs. the 2022 dataset (average of 29 tokens per speaker). In 1988, speakers who were born in the decades 1930's-1950's produced a rising-falling contour with peak aligned near the midpoint of the rhyme, but speakers who were born in the 1960's and 1970's produced contours with a late peak. This trend continues for the younger speakers recorded in 2022, with most speakers born in the 1980's and after producing a rising contour rather than a rising-falling contour.

Figure 8a) *Tone 4 f0 trajectory for each speaker from L-Thongkum (1988)'s dataset, which provided one token per speaker, drawn with geom_line function in ggplot2 (Wickham 2016);* ***8b)*** *Tone 4 f0 trajectory for each speaker from the current dataset, with multiple tokens fit to one polynomial curve per speaker with the geom_smooth function in ggplot2 (Wickham 2016)*

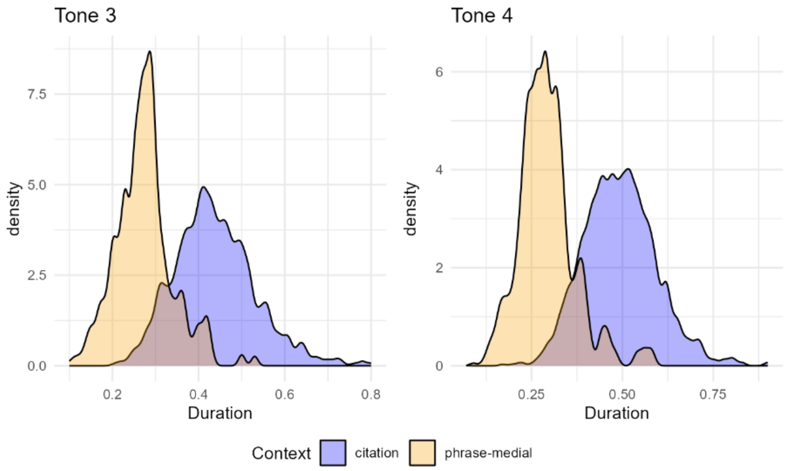


4.2 RQ2: Peak-delayed variant expands its distribution

In this section, we address Research Question 2: Does peak delay occur under conditions of preceding low tone and shorter duration, and does the peak-delayed variant expand its distribution across generations? We predict that peak delay will occur under the above conditions, which are both found in the phrase-medial context where targets were produced in a L_L frame with shorter duration. We also examine whether the peak-delayed variant changes its distribution over the course of several generations. If peak sliding is occurring as an ongoing change in T3 and T4, then we expect younger speakers to produce variants with a later peak alignment not only in conditions favorable to peak delay, but also in isolation, a context without a preceding low tone and often with longer duration than in connected speech.

In the present dataset, phrase-medial forms tend to be shorter in duration than citation forms (monosyllables pronounced in isolation), as can be seen in the distribution of Tone 3 and Tone 4 durations shown in Figure 9. We therefore cannot separately estimate the effects of context vs. duration on f0; however, teasing apart those effects is not necessary to answer RQ2. In the following GAMM model visualizations, we show short duration and phrase-medial context vs. long duration and citation context, to model conditions that favor or disfavor peak delay, respectively.

Figure 9: *Density plots of Tone 3 and Tone 4 durations by context. Citation forms tend to have longer durations than phrase-medial forms.*



4.2.1 Tone 3

Table 2 shows the results of model comparison between the full model and a model that excludes a particular predictor, in order to test whether that predictor has a significant effect on f0. For example, Decade of birth is represented in the full model as a smooth term and in interaction with timepoint (to model the effect on contour shape) and in an interaction with Duration (to model whether the effect of Duration on f0 depends on Decade of birth). Model comparison for Decade of Birth would then compare the full model with a model that excludes the above terms involving Decade of Birth. Results in Table 2 show that including predictors of Duration, Sex, Decade of birth, and Context significantly improves model fit. The interaction of Decade of birth and Duration was only found to be significant for phrase-medial context ($p<0.001$), suggesting that the effect of Duration on f0 changed across generations for the phrase-medial context but not for citation context.

Table 2: *Model comparison results for Tone 3*

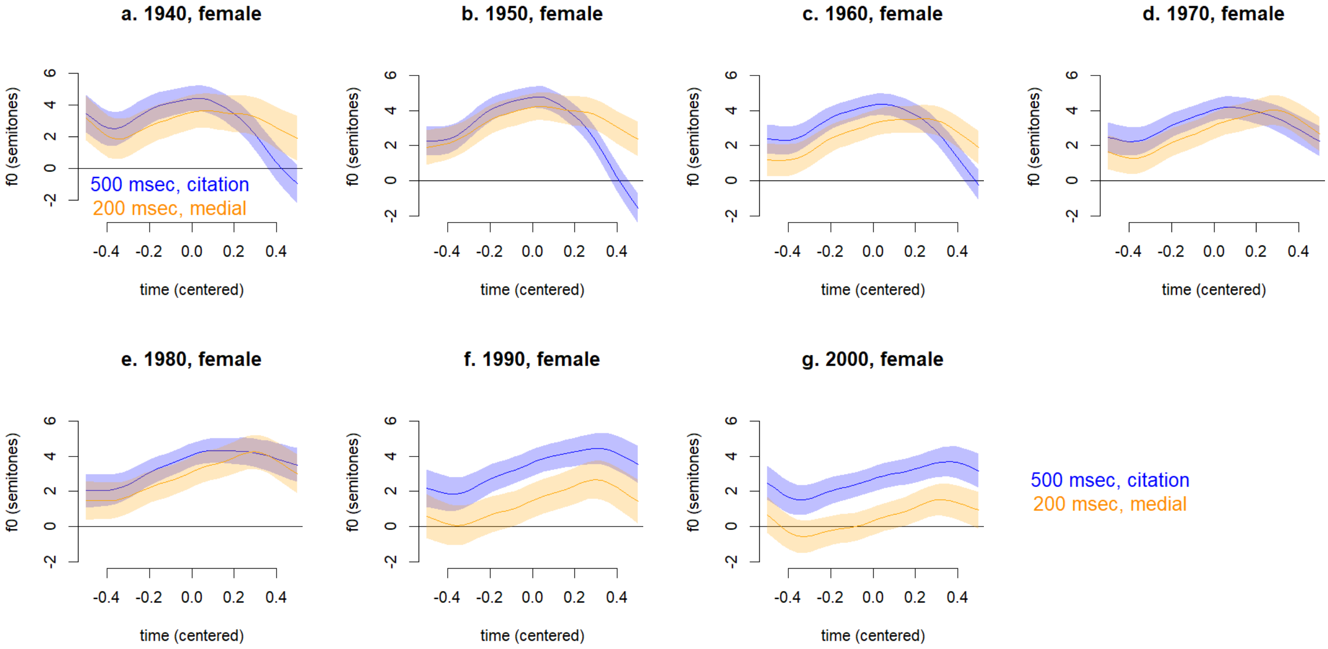
Comparison	χ^2	df	p
Duration	2237.123	32	<.0001***
Sex	1036.565	31	<.0001***
Decade of birth	918.202	32	<.0001***
Context	70.317	31	<.0001***

Figure 10 and Figure 11 display model predictions for T3 for female and male participants, respectively. The delay-favoring condition of short duration and phrase-medial context is shown in contrast to non-favoring condition of longer duration and isolation context. The panels are arranged by decade of birth.

Comparing across decades of birth, a gradual change in the shape of the contour from rising-falling to rising is apparent, a trend most strongly evident in female speakers (Figure 10). Furthermore, the younger a speaker is, the more likely that a peak-delayed variant will appear in both favoring and non-favoring conditions. For female speakers born in the 1940's and 1950's, distinct tonal alignments are seen in the two conditions: the delay favoring condition shows a later peak compared to the non-favoring condition. The later peak also co-occurs with reduction in the excursion of final fall, resulting in the delayed variant having more of a rising contour than a rising-falling one.

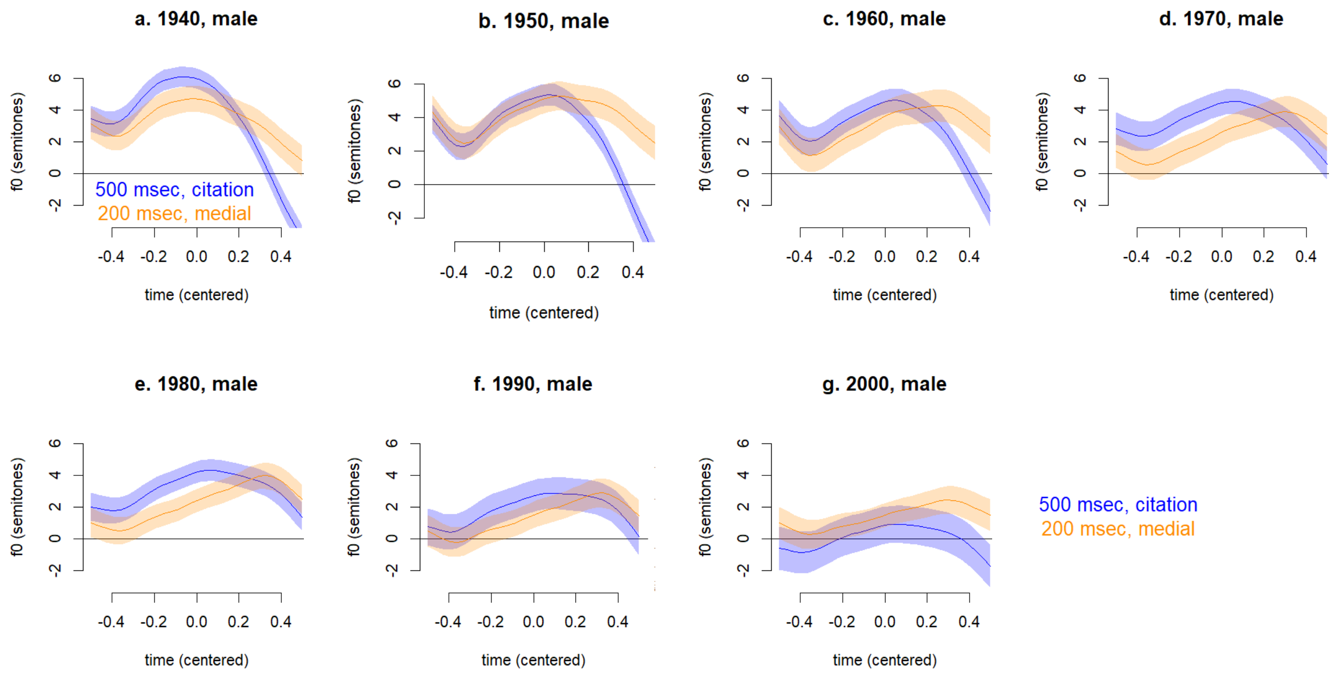
Female speakers born in the 1960's and 1970's also show a peak delayed variant in the favoring condition. In contrast with older speakers, however, the peak is located past the midpoint even in the non-favoring condition. In female speakers born in subsequent decades (1980's -2000's), the peak in both favoring and non-favoring conditions is located well past the midpoint, and the shape in both conditions is similar: a rising contour with only minimal final fall.

Figure 10: *Tone 3 GAMM fitted values for different durations and context (500 msec in isolation and 200 msec in medial position) for female speakers, arranged by decade of birth. Ribbon represents 95% confidence interval.*



For male speakers, as shown in Figure 11, the trends described above are also evident. Peak delay occurs in the favoring condition but not in the non-favoring condition for speakers born in the 1940's through 1950's, but for speakers born in the 1960's and later decades, later peaks and more of a rising contour shape are seen in both favoring and non-favoring conditions. Compared to females, male speakers born after 1960 more closely adhere to a rising-falling contour shape in non-favoring condition. Male speakers born in 1980's-2000's still show a recognizably rising-falling contour in the non-favoring condition, but the same cannot be said for female speakers.

Figure 11: *Tone 3 GAMM fitted values for different durations and context (500 msec in isolation and 200 msec in medial position) for male speakers, arranged by decade of birth. Ribbon represents 95% confidence interval.*



4.2.2 Tone 4

Table 3 shows model comparison results for the T4-T5 model. As discussion of potential T4-T5 merger is beyond this paper's scope, we focus here only on interpreting the T4 results relevant to the research questions. Similar to the T3 results, T4 model comparison indicates that Decade of birth, Duration, Sex, and Context had a significant effect on f0. In contrast with T3 results, the interaction of Decade of birth and Duration was only found to be significant for citation context (p<0.001), suggesting that the effect of Duration on f0 changed across generations for the citation context but not for the phrase-medial context.

Table 3: *Model comparison results*

Comparison	χ^2	df	p
Decade of birth	3845.705	48	<.0001***
Duration	1899.290	48	<.0001***
Sex	69.148	31	<.0001***
Context	63.766	31	<.0001***

Figure 12 and Figure 13 present model predictions for T4 for female and male participants, in the same arrangement as Figures 10 and 11. In contrast with the result for T3, female speakers born in the 1940's and 1950's do not produce distinct shapes according to condition (Figure 12): both conditions show a rising-falling contour. However, female speakers born in the 1960's and 1970's produce late peak rising-falling variants in the non-favoring condition and rising or falling-rising variants in the favoring condition. For female speakers born in the 1980's and later decades, the shape in both conditions is low rising with a more distinctly "scoopy" rise compared to older speakers' contours.

Figure 12: *Tone 4 GAMM fitted values for different durations and context (500 msec in isolation and 200 msec in medial position) for female speakers, arranged by decade of birth. Ribbon represents 95% confidence interval.*

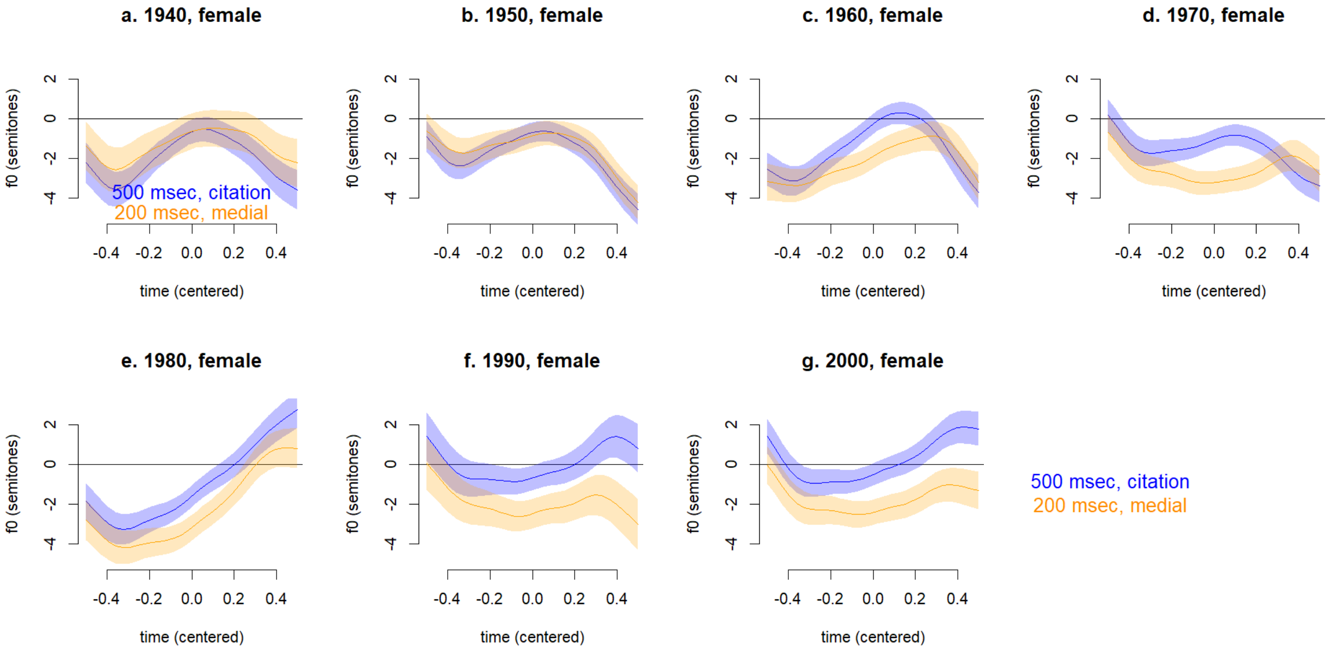
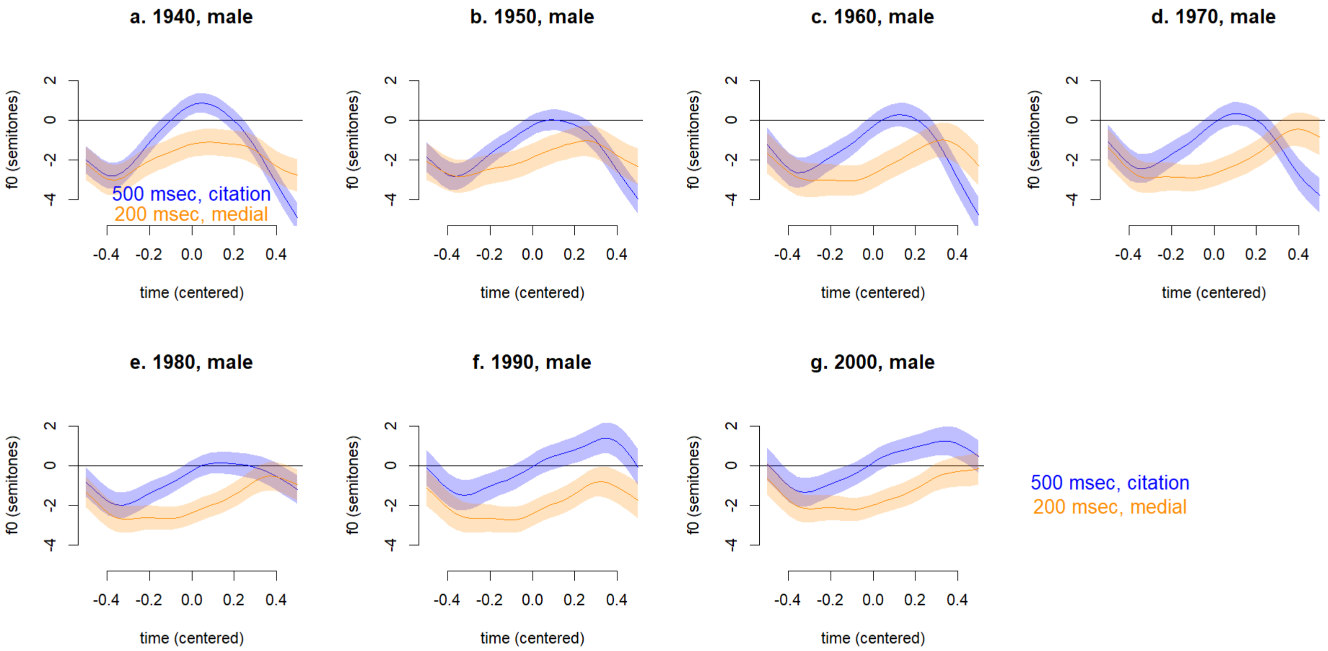


Figure 13 shows that male speakers born in the 1940's and 1950's produce distinct shapes by condition: later peaks and less f0 excursion (thus, more of a rising contour) are seen in the favoring condition. Males speakers born in the 1960's-1980's maintain this conditioned variation, but those born in the 1990's and 2000's show similar contour shapes (rising) in both conditions. That is, younger speakers produce a peak-delayed variant in both favoring and non-favoring conditions, a pattern that is also seen in Tone 3 results.

Figure 13: *Tone 4 GAMM fitted values for different durations and context (500 msec in isolation and 200 msec in medial position) for male speakers, arranged by decade of birth. Ribbon represents 95% confidence interval.*



5 Discussion and Conclusion

In response to RQ1, we find evidence of monotonic change in real time from rising-falling to rising contours in Tones 3 and 4 over the 34 years from 1988 to 2022. Younger speakers increasingly produce rising variants. This trend study's results confirm L-Thongkum's (1997) predictions based on data collected in 1988. Female speakers appear to be leading the change from rising-falling to rising contours in Iu Mien tones, producing more advanced variants compared to male speakers of the same generation. This result is in line with well-established gender patterns in sound change in which women lead changes toward incoming prestige variants (Labov 1990).

Turning to RQ2, peak delay occurred as predicted under conditions of a preceding low tone and shorter duration. Critically, the peak-delayed variants expanded their distribution to non-favoring contexts (longer duration, citation form) across generations. Younger speakers produced the peak-delayed variants not just in favoring contexts but also in non-favoring contexts. For both tones, an age vector associating peak-delayed variants with younger speakers has emerged. The increasing use of the peak-delayed rising variant among generations born after 1970 most likely reflects "momentum-sensitive learning", in which younger speakers align their pronunciation with the perceived age vector, cf. Bermúdez-Otero (2020).

L-Thongkum's (1997) explanation that contact with Standard Thai drove these changes is supported by the timeline, as dramatic increases in Thai proficiency occurred after 1978 when a primary school was built in the village (p. 155). The observed changes accelerated amongst generations born after 1970 who gained greater Thai proficiency through education. The present study cannot definitively tease apart the contributions of language contact versus peak delay. However, the present study's results show that peak delay generates rising variants from rising-falling tones, providing a plausible phonetic precursor for the changes, even if increasing use of Thai may have triggered the changes to initiate at that particular time.

Trend studies are valuable in establishing the direction of change in real time. However, few such studies exist on tone change. The current study contributes to understanding of tone change through documenting a case study of rising-falling > rising across the course of 34 years (1988-2022). The present study's results align with apparent-time case studies of other Asian languages showing the same pathway of change: Bangkok Thai (Pittayaporn 2018), Wuxi Wu (Zhang 2014), Jining Mandarin (Wang 2006), and Lanzhou Mandarin (Zhang 2012), as reviewed in Yang & Xu (2019). In contrast, rising > rising-falling has not been reported in any ongoing tone change study to our knowledge, which implies that rising-falling > rising may be a unidirectional trend. Rising is not expected to develop into rising-falling, because peaks rarely slide leftwards in connected speech. This does not mean that simple contour tones will never develop into complex contour tones. Indeed, preliminary evidence suggests that rising-falling may develop from high falling (Yang, Kirby & Dockum 2023).

This study helps address the critical gap in documentation of tone change in real time. Further research on a larger sample of Iu Mien varieties is needed to determine if the changes are occurring in other dialects. Investigating the social factors influencing individual variation would also shed light on propagation of the changes through the community. By establishing a phonetic motivation for the (potentially) unidirectional pathway rising-falling > rising, this study moves us closer to solving the constraints problem for tone change.

Acknowledgements

The authors thank Mrs. Tamami Arisawa, Lawyer Phitphichet Phanphisutthichon and Mrs Phichayanan Phanphisutthichon, the President of Mae Yao Local Administration Mr. Songkran Aphinyawisan, and the Iu Mien communities of Huai Mae Sai and Huai Kaeo for their assistance with data collection and community engagement. This research was funded by a grant from SIL Mainland Southeast Asia Group's Asian Minority Language Research Committee.

Supplemental material

Data and R code used to analyze and produce figures are available for viewing and downloading at https://doi.org/10.5281/zenodo.10597579

Disclosure statement on use of AI

During the preparation of this work the authors used ClaudeAI in order to improve readability and language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Abramson, Arthur S. 1979. The coarticulation of tones: An acoustic study of Thai. In Theraphan L. Thongkum, P. Kullavanijaya, V. Panupong & K. Tingsabadh (eds.), *Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugenie J.A. Henderson*, 1–9. Bangkok: Chulalongkorn University Press.

Barnes, Jonathan, Nanette Veilleux, Alejna Brugos & Stefanie Shattuck-Hufnagel. 2012. Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology* 3(2). 337–383. https://doi.org/10.1515/lp-2012-0017

Bermúdez-Otero, Ricardo. 2020. The initiation and incrementation of sound change: Community-oriented momentum-sensitive learning. *Glossa: A Journal of General Linguistics* 5.1: 121. 1–32. https://doi.org/10.5334/gjgl.627

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Boersma, Paul & David Weenick. 2023. Praat: doing phonetics by computer. https://www.praat.org

Bradley, Cornelius Beach. 1911. Graphic analysis of the tone-accents of the Siamese language. *Journal of the American Oriental Society*. American Oriental Society 31.3:282–289. https://doi.org/10.2307/3087645

Burroni, Francesco. 2023. *Dynamics of F0 planning and production: Contextual and rate effects on Thai tone gestures*. Ithaca, N.Y.: Cornell University Ph.D. dissertation.

Bybee, Joan. 1999. Usage-based phonology. In Michael Darnell, Edith A. Moravcsik, Michael Noonan, Frederick J. Newmeyer & Kathleen Wheatley (eds.), *Functionalism and Formalism in Linguistics: Volume I: General papers*, 211–242. John Benjamins Publishing.

Bybee, Joan L. 2012. Patterns of lexical diffusion and articulatory motivation for sound change. In Maria-Josep Solé & Daniel Recasens (eds.), *Current Issues in Linguistic Theory*, vol. 323, 211–234. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/cilt.323.16byb

Campbell, Eric W. 2021. Why is tone change still poorly understood, and how might documentation of less-studied tone languages help? In Patience Epps, Danny Law & Na'ama Pat-El (eds.), *Historical linguistics and endangered languages: Exploring diversity in language change*, 15–40. New York: Routledge.

Chitoran, Ioana. 2012. The nature of historical change. In Abigail C. Cohn, Cécile Fougeron & Marie K. Huffman (eds.), *The Oxford Handbook of Laboratory Phonology*, 311–321. Oxford University Press.

Chuang, Yu-Ying, Janice Fon & R Harald Baayen. 2021. Analyzing phonetic data with generalized additive mixed models. In Martin J. Ball (ed.), *Manual of Clinical Phonetics*, 108–138. Milton, U.K.: Taylor & Francis Group. http://ebookcentral.proquest.com/lib/uhm/detail.action?docID=6481866 (2 December, 2022).

Eberhard, David M., Gary Simons & Charles D. Fennig (eds.). 2023. *Ethnologue: Languages of the world*. 25th edn. SIL International. https://www.ethnologue.com/. (9 September, 2022).

Garrett, Andrew & Keith Johnson. 2013. Phonetic bias in sound change. In Alan C. L. Yu (ed.), *Origins of sound change: Approaches to phonologization*, 51–97. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573745.003.0003

Harrington, Jonathan, Felicitas Kleber, Ulrich Reubold, Florian Schiel & Mary Stevens. 2019. The phonetic basis of the origin and spread of sound change. In William F. Katz & Peter F. Assmann (eds.), *The Routledge Handbook of Phonetics*, 401–426. 1st edn. New York, NY: Routledge. https://doi.org/10.4324/9780429056253-15

Karlin, Robin. 2018. *Towards an articulatory model of tone: a cross-linguistic investigation*. Ithaca, N.Y.: Cornell University PhD dissertation. https://ecommons.cornell.edu/handle/1813/64937 (7 March, 2023).

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2.2:205–254. https://doi.org/10.1017/S0954394500000338

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Vol. 1. Oxford, UK ; Cambridge, MA: Blackwell.

L-Thongkum, Theraphan. 1988. *Acoustical measurements of tones and pitch ranges in the Mien-Yao dialect of Huai Mae Sai village, Muang district, Chiangrai province (In Thai)*. Unpublished research report. Bangkok: Chulalongkorn University.

L-Thongkum, Theraphan. 1997. Tone change and language contact: A case study of Mien-Yao and Thai. In A. S. Abramson (ed.), *Southeast Asian Linguistics Studies In Honor of Vichin Panupong*, 153–160. Bangkok: Chulalongkorn University Press.

Ohala, John J. 2003. Phonetics and historical phonology. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 669–686. Malden, MA: Blackwell.

Pittayaporn, Pittayawat. 2018. Phonetic and systemic biases in tonal contour changes in Bangkok Thai. In Haruo Kubozono & Mikio Giriko (eds.), *Tonal change and neutralization* (Phonology and Phonetics), vol. 27, 249–278. Berlin: Mouton de Gruyter.

Prasertsud, Sumeth. 2017. *Iu Mien - English - Thai Dictionary*. Pongsawatgarnpim, Ltd.

Purnell, Herbert. 2012. *Iu-Mienh-English Dictionary: With Cultural Notes*. Chiang Mai: Silkworm Books. https://digitalcommons.biola.edu/faculty-books/60

R Core Team. 2023. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org

Ratliff, Martha. 2010. *Hmong-Mien language history*. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University. https://doi.org/10.15144/PL-613

Rij, Jacolien van, Martijn Wieling, R. Harald Baayen & Hedderik van Rijn. 2022. Itsadug: interpreting time series and autocorrelated data using gamms.

Sóskuthy, Márton. 2017. Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. arXiv. https://doi.org/10.48550/arXiv.1703.05339

Sóskuthy, Márton. 2021. Evaluating generalised additive mixed modelling strategies for dynamic speech analysis. *Journal of Phonetics* 84:101017. https://doi.org/10.1016/j.wocn.2020.101017

Sóskuthy, Márton, Paul Foulkes, Vincent Hughes & Bill Haddican. 2018. Changing words and sounds: the roles of different cognitive units in sound change. *Topics in Cognitive Science* 10.1:787–802. https://doi.org/10.1111/tops.12346

Thomas, Erik K. 2019. Innovation in sociophonetics. In William F. Katz & Peter F. Assmann (eds.), *The Routledge handbook of phonetics*, 448–472. New York: Routledge.

Wang, Yufeng. 2006. Cong Neimenggu Jining fangyan xinlao liangpai shengdiao xitong de bianyi kan shengdiao yanbian (从内蒙古集宁方言新老两派声调系统的变异看声调的演变) [Examining tonal evolution through the tone system variation of new and old Jining dialect of Inner Mongolia]. *Zhongguo Yuwen* (1). 41–44.

Weinreich, Uriel, William Labov & Marvin I Herzog. 1968. Empirical foundations for a theory of language change. In Lehmann, Winifred & Malkeil, Yakov (eds.), *Directions for Historical Linguistics*, 95–195. Austin: University of Texas Press.

Wickham, Hadley. 2016. *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wieling, Martijn. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics* 70:86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. 2nd edn. Chapman and Hall/CRC.

Xu, Yi. 2001. Fundamental frequency peak delay in Mandarin. *Phonetica* 58.1–2:26–52. https://doi.org/10.1159/000028487

Xu, Yi. 2006. Tone in connected discourse. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, vol. 12, 742–750. Oxford: Elsevier.

Xu, Yi & Xuejing Sun. 2002. Maximum speed of pitch change and how it may relate to speech. *The Journal of the Acoustical Society of America* 111.3:1399–1413. https://doi.org/10.1121/1.1445789

Yang, Cathryn, James P. Kirby & Rikker Dockum. 2023. The diachronic origins of complex contour in Tai. Paper presented at the 56th International Conference on Sino-Tibetan Languages and Linguistics, Chulalongkorn University, Bangkok.

Yang, Cathryn & Yi Xu. 2019. Crosslinguistic trends in tone change: A review of tone change studies in East and Southeast Asia. *Diachronica* 36.3:417–459.

Yu, Alan C. L. 2021. Toward an individual-difference perspective on phonologization. *Glossa: a journal of general linguistics*. Ubiquity Press 6.1:14. https://doi.org/10.5334/gjgl.661

Zhang, Guanyu. 2012. Lanzhou danziyin shengdiao geju de tongji fenxi (兰州话单字音声调格局的统计分析) [Statistical analysis of Lanzhou citation tone patterns]. *现代语文(语言研究版) [Modern Chinese (Language Research)]* 2012(3):12–18.

Zhang, Jingwei. 2014. *A sociophonetic study on tonal variation of the Wúxī and Shànghǎi dialects*. Utrecht: LOT.

Zsiga, Elizabeth & Rattima Nitisaroj. 2007. Tone features, tone perception, and peak alignment in Thai. *Language and Speech* 50.3:343–383.

Reviewed: Received 2 February 2024, revised text accepted 10 April 2024, published 1 May 2024
Editors: Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Sigrid Lew, Dr Mathias Jenny, Dr Kenneth Van Bik