Change and stability in the tonal contours of King Rama IX of Thailand, 1959-1997

Cathryn Yang1, Pittayawat Pittayaporn2, James Kirby3, Sujinat Jitwiriyanont2

1SIL International and Payap University, Thailand
2Southeast Asian Linguistics Research Unit, Faculty of Arts, Chulalongkorn University, Thailand
3Institute of Phonetics and Speech Processing, Ludwig Maximilian University of Munich, Germany
cathryn_yang@sil.org, pittiyawat.p@chula.ac.th, jkirby@phonetik.uni-muenchen.de, sujinat.j@chula.ac.th

Abstract

Studies of change across the lifespan have enabled a more nuanced understanding of segmental sound change, but this method has hardly ever been used in tone change research. This study, one of the first longitudinal studies of tone, investigates the tonal contours of King Rama IX of Thailand across a 38-year period (1959-1997). During the 20th century, Thai Tone 3 (“Falling”) changed from mid-falling to high-falling. The study examines whether King Rama IX’s T3 pronunciation changed during the period of investigation, and if so, whether the change was in the same direction as the community trend. The data consist of 16 recordings of the King’s commencement speeches at Chulalongkorn University. F0 was extracted from 1317 T3 tokens and fitted to a cubic polynomial function. Linear mixed model results and post-hoc analysis of estimated marginal means indicate a period of change followed by stability: T3’s F0 onset raised and the slope became steeper between the 1960’s and the 1980’s (in line with community trends), but not between the 1980’s and 1990’s. These findings align with previous research showing that adults have the potential to change in the post-adolescent period, but they also may resist change when that aligns with their identity.

Index Terms: Bangkok Thai, tone change, lifespan change

1. Introduction

Although it is clear that tone systems continue to evolve after tonogenesis, the mechanisms and constraints that govern phonetic tone change remain poorly understood [1]. Panel studies (i.e., those comparing data from the same individual(s) at different time periods) have enabled a more nuanced understanding of segmental sound change, but this method, to the best of our knowledge, has only been used once in tone change research [2]. In this study, we investigate tone pronunciation across the lifespan through an acoustic analysis of the tone contours of King Rama IX of Thailand over a 38-year period (1959-1997). Standard (Bangkok) Thai has five tones (see Figure 2), labeled “T1-Mid”, “T2-Low”, “T3-Falling”, “T4-High”, and “T5-Rising”. During the 20th century, Bangkok Thai tone contours changed noticeably in the community, with T3 showing a dramatic change from mid falling to high falling [1]. The purpose of this study is to investigate whether Rama IX’s T3 production shows evidence of change during the period of investigation, and if so, whether the change was in line with community trends.

As reviewed in [3] and [4], stable pronunciation of segments during adulthood is the most common finding of lifespan studies, but there are also cases of change towards and against community trends, e.g. [4, 5]. In a longitudinal study of tone in a single speaker of Hong Kong Cantonese, [2] found that tone pronunciation became more conservative over the course of three decades, as the speaker became more involved with conservative political and language ideology. This study shows that tone production can change well into adulthood and may not always conform to broader community trends.

The subject of the current study, King Rama IX (1927-2016), moved back to Thailand at age 25 after spending much of his childhood and youth in Switzerland [6]. Panel studies such as [7] and [8] have shown that, when an individual relocates to a region with different dialect features, they may adapt at least partially to the norms of their new community. The case of Rama IX is somewhat different from the above case studies, as Rama IX was a native speaker of Bangkok Thai from birth. However, as he grew up abroad, his pronunciation may not have been identical to his age cohort in Bangkok, particularly for variables that were undergoing change in Bangkok at the time, cf. [9]. This scenario, in which a native speaker returns to their homeland after growing up abroad, has not yet been examined in lifespan change studies, but we expect it to share some similarity with cases in which speakers move to a new dialect region.

2. Evolution of Bangkok Thai T3 contours in the 20th century

Several related changes to T3 occurred during the 20th century, illustrated by comparing Figure 1 [10], a tone plot from the early 20th century, to Figure 2 [11] from the late 20th century. The first change to occur, as discussed in [1], was F0 onset raising from mid (as in Figure 1) to high (as in Figure 2). F0 onset raising of a mid falling tone appears to be a phonetically motivated tone change, as suggested by the existence of diachronic phonetic precursors and the frequency of this change occurring independently in non-related languages. Preceding segments that are voiceless and preceding tones that have high F0 offsets both have raising effects on following F0 onsets [1, 12]. In a recent review of tone change studies, mid falling > high falling ranked in the top five most frequently reported changes across a database of 45 Asian languages from various language families [13].

This gradual change must have begun in Bangkok after 1911 but before the 1950’s. Tone plots in [14] based on recordings from the late 1950’s show a T3 onset that is already slightly above mid. The speech analyzed in [14] was spoken by a Bangkok male born in 1930, who thus belongs to the same
T3 F0 onset raising set the stage for two subsequent phonetic changes: peak sliding and the development of a truncated variant in non-final position. Once the F0 onset became high, timing constraints on the realization of a high tone target came into play, resulting in a gradual rightward shift in tonal alignment. In connected speech, reaching a high tone target often entails a period of rising F0, and rises require a longer time to produce than level or falling contours [15]. In free speech, there are thus many instances of peak delay, in which a high target is realized late in the syllable or even in the onset of the following syllable [16]. Diachronically, this variation sometimes leads to peak sliding [1]: the peak shifts later in the syllable, resulting in a rising-falling contour (as in Figure 2). As the peak slides rightward, the timing of the following fall slides rightward as well, which results in truncation of the fall in non-final position due to an inadequate amount of time (see Figure 3). This development was first observed in young speakers in the early 1990’s and documented in [17].

3. Method

3.1. Subject

In this section we discuss biographical details of King Rama IX that are pertinent to the study. Rama IX’s family language was Thai, and he lived in Bangkok during his early childhood. However, he spent a large portion of his childhood and youth (age 6-21) in Switzerland. [6]. Growing up abroad, Rama IX did not have frequent interaction with his age cohort in Bangkok. Daily interactions in Thai were limited to his mother, siblings, and Thai tutors [6]. His mother (b. 1900) and the Thai tutors were most likely speaking a form of Thai that followed older pronunciation norms than that of Rama IX’s age cohort in Bangkok. As a diasporic youth, Rama IX probably modeled his speech on this older generation rather than his age cohort in Bangkok, as suggested by the findings in [9]. Rama IX returned to live permanently in Thailand in 1951 at the age of 25 [6].

Throughout his long reign, Rama IX evinced a language ideology that emphasized adherence to a national standard. On numerous occasions, he spoke on the importance of maintaining national language standards. At the same time, Rama IX’s accent was perceived by some to be provincial, a perception that Rama IX discussed publicly in a 1992 speech [18]. In the speech, he explained that the pronunciation used by people from outside the capital reflected an older and therefore more correct form of the language, in contrast to the changing speech of the capital. He specifically noted the contour changes to T3 and T4: “The pitch keeps going higher. Tone 3 has become Tone 4. Tone 4 has become Tone 5. These changes make [present-day speech] sound unusual.” The likening of T3 contour to that of T4, which traditionally had a rising-falling contour, probably referred to T3’s relatively new rising-falling variant. His observation implies that T3’s rising-falling variant had become salient by the early 1990’s.

3.2. Recording collection


Transcripts and video recordings (audio only before 1982) are available at http://www.kingrama9.chula.ac.th/. Video recordings were converted to audio (WAV) format with a 44,100 Hz sampling rate. Mean duration of the recordings was 4 minutes 25 seconds.
The setting, content and style is consistent across the recordings. In each recording, Rama IX is reading a speech in a formal register of Standard Thai (i.e., the national standard based on the pronunciation of Bangkok Thai elite). Tonal differences are not usually noted in comparisons of different registers in Thai [19]. However, formal speech is often characterized by a slower speech rate compared to a normal speech rate, and [20] found that, in Thai, larger F0 excursions were found in slow speech compared to fast speech. Rama IX’s speech rate is slower than normal conversation, having a mean of 2.15 syllables/second across the recordings. For the purposes of this study, however, the critical question is whether his speech rate remains consistent across the period under investigation, which Table 1 clearly shows is the case.

Table 1: Length and speech rate (syllables per second) of each recording

<table>
<thead>
<tr>
<th>Decade</th>
<th>rec. minutes</th>
<th>syll/sec</th>
<th>rec. minutes</th>
<th>syll/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>1959</td>
<td>5.86</td>
<td>2.11</td>
<td>1993.1</td>
<td>4.43</td>
</tr>
<tr>
<td>1969</td>
<td>2.87</td>
<td>2.49</td>
<td>1993.2</td>
<td>5.36</td>
</tr>
<tr>
<td>1982</td>
<td>4.53</td>
<td>2.07</td>
<td>1995.1</td>
<td>4.13</td>
</tr>
<tr>
<td>1987_1</td>
<td>4.84</td>
<td>2.22</td>
<td>1995.2</td>
<td>3.78</td>
</tr>
<tr>
<td>1987_2</td>
<td>4.49</td>
<td>2.33</td>
<td>1995.3</td>
<td>4.24</td>
</tr>
<tr>
<td>1987_3</td>
<td>5.47</td>
<td>2.14</td>
<td>1996.1</td>
<td>3.84</td>
</tr>
<tr>
<td>1989</td>
<td>4.59</td>
<td>2.19</td>
<td>1997.1</td>
<td>3.67</td>
</tr>
<tr>
<td>1991</td>
<td>3.92</td>
<td>2.10</td>
<td>1997.2</td>
<td>4.68</td>
</tr>
</tbody>
</table>

3.3. Data processing

Transcripts were manually corrected in OCTRA [21], including marking disfluencies and pauses. The Munich Automatic Segmentation System [22] was used to perform automatic forced alignment for segmentation and labeling. Tone was automatically labeled according to the citation tone, as there is no tone sandhi in Thai. Segmentation and labeling were manually corrected by native-speaking research assistants.

Data processing was carried out in the EMU Speech Database Management System [23]. F0 in syllables ending with a vowel or a sonorant was calculated at 5 msec intervals using the Praat pitch tracker [24]. Measurement of F0 began at the onset of the vowel and continued to the end of the syllable.

The following types of tokens were identified and filtered out: tokens without a tone label (-7), tokens without an initial consonant (-7), tokens with no F0 values (-40), tokens with F0 onset greater than 10 semitones above mean F0 (-13), tokens with less than 50 msec of voicing (-362), and tokens that showed a sudden F0 jump of over 60 Hz (-138). This filtering resulted in a loss of 567 tokens (9% of the total) and a final count of 1317 T3 tokens. Table 2 provides token counts by decade for each tone.

Table 2: Token counts by decade for each tone

<table>
<thead>
<tr>
<th>Decade</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1959-1969</td>
<td>310</td>
<td>40</td>
<td>145</td>
<td>42</td>
<td>85</td>
<td>622</td>
</tr>
<tr>
<td>1982-1989</td>
<td>944</td>
<td>138</td>
<td>498</td>
<td>130</td>
<td>272</td>
<td>1982</td>
</tr>
<tr>
<td>1991-1997</td>
<td>1535</td>
<td>196</td>
<td>674</td>
<td>205</td>
<td>351</td>
<td>2961</td>
</tr>
</tbody>
</table>

F0 was time-normalized and scaled by converting Hz to semitones relative to year-specific F0 means. Year-specific mean was used as the baseline because mean F0 usually lowers with increasing age due to physiological factors (see Figure 4). F0 track data were fitted to a cubic polynomial function to represent four parameters of tone contour shape: intercept (F0 onset), initial slope, curvature, and S-shape. Model comparison via Anova was used to confirm that a cubic function improved the fit significantly over a quadratic function, following [25], [26].

3.4. Analysis

Separate linear mixed effect models were fitted for F0 onset, slope, and curvature using the lme4 package in R [27]. Because there was a gap in the data from 1970-1981, we grouped the data into periods corresponding roughly to decade: 1959-1969 “the 1960’s”, 1982-1989 “the 1980’s”, and 1991-1997, “the 1990’s”. Fixed factors included Period (our critical term of interest) and phonetic factors predicted to affect the dependent variable: Duration (centered), Preceding tonal context (six levels: Tones 1-5 and silence), Following tonal context (same levels as Preceding context), and Preceding segment (three levels: sonorant, voiced obstruent, voiceless obstruent). All categorical predictors were Helmert coded. Interactions were not included as they did not significantly improve model fit. The model includes Syllable as a random intercept and by-syllable random slopes for Period.

4. Results

4.1. Biological aging effects on F0

Figure 4 shows that F0 mean decreases over time during the age range of 32-70 years, reflecting a typical biological effect of increasing age on F0. This result is consistent with results from other longitudinal studies of male speakers such as [7]. Note that biological effects on F0 are controlled for in the analysis by using year-specific mean as the baseline for converting F0 to semitones.

![Figure 4: Boxplot of F0 range by year](image)

4.2. T3 onset

Figure 5 compares Rama IX’s tone system in the three time periods (the 1960’s, 1980’s and 1990’s). One of the most striking differences over time is seen in T3, highlighted in red. T3 F0 onset increases, and the slope becomes more steeply falling between the first two periods. There is little difference between the second and third time periods.

Modeling results corroborate the trends apparent in Figure 5. We report model-based pairwise comparison of the estimated marginal means [28]. For F0 onset, there was a 0.77 semitone increase between the 1960’s and the 1980’s (SE = 0.20, df = 20.1, t = 3.93, p = 0.002) and a 0.81 increase between the 1960’s and the 1990’s (SE = 0.21, df = 25.6, t = 3.9, p = 0.002). The difference between the 1980’s and 1990’s, however, was not significant. Unsurprisingly, there were phonetic contextual
effects as well: a preceding Tone 3 had a lowering effect, as did longer duration, while a preceding Tone 4 or Tone 5, a preceding voiceless obstruent, or a following Tone 1 or Tone 2 had a raising effect.

4.3. T3 slope

Figure 6 shows that the slope becomes steeper between the 1960’s and 1980’s, a consequence of the fact that Rama IX’s T3 onset raises considerably, but the offset does not change much. There was a 1.7 increase in the falling slope between the 1960’s and the 1980’s (SE = 0.31, df = 34.9, t = 5.5, p < 0.0001), followed by a 0.44 decrease (i.e., flattening) between the 1980’s and 1990’s (SE = 0.17, df = 30.3, t = -2.5, p = 0.04). Overall, there was a 1.2 increase between the 1960’s and the 1990’s (SE = 0.26, df = 30.3, t = 4.68, p = 0.0002). As for phonetic effects, there was a flattening of the slope after a voiced obstruent or before a T3 and an increase in slope with longer duration.

4.4. T3 curvature

As discussed in Section 2, after T3 onset raised to high, the peak slid later in the syllable, creating a rising-falling variant in the community (see Figure 2). If Rama IX had participated in this change, we would expect T3 curvature to change from near zero (representing a linear contour) to negative (denoting a rising-falling contour). However, there was no significant difference in curvature between any periods. The evidence is insufficient to support a claim that T3 contour moved toward a rising-falling shape during the period of investigation.

5. Discussion

The study’s results indicate that Rama IX’s T3 pronunciation did change, though in a non-linear fashion. His T3 F0 onset raised, and the slope became steeper between the 1960’s and the 1980’s, but later remained stable. The direction of the observed changes is in line with broader community trends that saw T3 moving from mid falling to high falling during the 20th century. This result is comparable to [29], a study that found, during a specific period, Queen Elizabeth’s pronunciation of certain vowels moved toward a norm that was spreading through the community at that time. One notable difference is that Rama IX grew up abroad and may have modeled his speech on an older generation rather than his Bangkok age cohort. The results suggest that, after he returned to Bangkok at age 25, he adapted to the norms of his age cohort. Some degree of post-adolescent adaptation has also been found in other studies such as [7] and [8].

Rama IX’s T3 changed significantly during middle adulthood, but then remained stable in late adulthood. He does not follow the secondary changes of developing a rising-falling contour or a more truncated variant in non-final position. This result indicates that, once adapted to the norms of his age cohort, he adhered to those norms in later years and resisted ongoing changes that were led by younger generations. His resistance to the later changes aligned with his language ideology. In this sense, the study’s results show similarities to [2], which found that Hong Kong politician C.Y. Leung resisted tone merger as he adopted a language ideology that promoted an older pronunciation norm.

6. Conclusions

The first longitudinal study of Thai tone shows King Rama IX undergoing a period of change after a major life transition, followed by a period of stability. This finding is congruent with two of the main findings of segmental longitudinal studies, namely, stability (the most common finding) and lifespan change in the direction of community trends (the second most common finding) [3].

7. Acknowledgements

We wish to thank Pawitsapak Akarajaradwong, Kantika Apinyovichien, Pimchip Kochaiyaphum, Peeranit Nittuyanont, Phongbhorn Prayongdravya, Kedsaraporn Ruechai, Thitivut Sommool, Uracha Tatiyanunt, and Chanakan Wittayasakpan for their work on data preparation, which was made possible by the CLARIN infrastructure. Chulalongkorn International Communication Center helped in acquiring the speech materials. This project was funded by the Ratchadapisek Somphot Fund as part of the annual support for the Southeast Asian Linguistics Research Unit, Faculty of Arts, Chulalongkorn University, and by ERC EVOTONE project (grant no. 758605) to J. Kirby.

8. References
