The interaction between language usage and acoustic correlates of the Kuy register distinction

Raksit T. Lau-Preechathammarach
Department of Linguistics, University of California, Berkeley, USA
raksit@berkeley.edu
Abstract

Contact is often cited as an explanation for the convergence of areal features and has been proposed as an explanation for the emergence of tonal languages in Mainland Southeast Asia. The current production study probes this hypothesis by exploring the relationship between tonal language usage and the acoustic correlates of the register distinction in Kuy, a Katuic language, as spoken in a quadrilingual (Kuy, Thai, Lao, Khmer) Kuy community in Northeast Thailand. The results demonstrate greater persistence of fundamental frequency (f0) differences over the course of the vowel alongside more tonal language experience for male speakers; however, analysis of individual differences finds that $H_1^*-H_2^*$, a correlate of voice quality, is the primary cue for male speakers with greater tonal language experience. For female speakers, a trading relation is found between f0 and other register cues alongside tonal language experience at both the group and individual levels. These findings provide evidence for a model by which contact may serve to enhance existing, non-primary cues in a phonological contrast by shifting cue distributions, thereby increasing the likelihood that these cues will come to be perceived as prominent and phonologized.

Keywords: sound change; cue weighting; language contact; bilingualism; register; tonogenesis

1 Background

Areal linguistic features that are shared across genetically unrelated languages are generally understood to result from extended language contact. The literature is rife with examples of contact effects, although in many cases, it is difficult to tease apart the role of contact in language change from a structural predisposition for the given change and from mere coincidence. Closely tied to work on language contact is the experimental literature on bilinguals’ usage of L1 phonetic cues in the production and perception of contrasts in an L2 language and of the effects of L2 language exposure and usage on L1 phonological categories. The current study contributes to the intersection of this literature by exploring how variation may be structured by bilinguals’ language usage and experience, setting the stage for larger-scale contact effects. The phenomenon of interest is the register contrast in a variety of Kuy (Katuic; Austroasiatic) spoken in Thailand at the borders of Cambodia and Laos. As in other register languages, the Kuy register distinction is characterized by a cluster of cues, including fundamental frequency (f0), amplitude, voice quality (or, more narrowly, phonation), and vowel quality. Speakers, however, differ in the extents to which each cue manifests in the register distinction; the object of study is the relationship between the distribution of cues and language experience and usage patterns, particularly at the individual level.

Kuy populations in Thailand live in a historically quadrilingual society, but the continual increasing centralization of Thailand has pushed language usage to trend largely towards Kuy-Thai bilingualism. As standardization pressures lead Kuy speakers to increasingly shift to the national language, Thai, we expect changes involving cues that are shared between these languages, such as f0, which is the primary cue for the tonal contrast in Thai but only one of several acoustic cues for the register contrast in Kuy. The production data reported in this study show a clear shift in cue prominence in the register contrast that correlates with language experience. Most notably, f0 differences between the registers show a marked increase alongside decreased differences in other register cues for female speakers who use Thai/Lao more at both the group and individual level. This increase occurs in concert with the weakening of voice quality cues. For male speakers however, the evidence is mixed—group differences show an increasing importance of f0 cues, but individual differences do not. There is a clear shift in male speakers towards the
usage of $H_1^* - H_2^*$, a robust acoustic correlate of voice quality crosslinguistically. I argue that these results show how the rearrangement of cue weights at the individual level can be shaped by patterns of bi- and multilingualism and accumulate to bring about societal-level sound changes that may be understood as language contact effects.

1.1 Cues, Contrast, and Contact

Tone is one of the most well-known areal features of Mainland Southeast Asia (MSEA), an area that comprises five large families: Sino-Tibetan, Kra-Dai, Hmong-Mien, Austro-Asiatic, and Austronesian (Henderson, 1965; Matisoff, 2001; Enfield, 2005, 2011). In a sample of 186 MSEA languages, Kirby & Brunelle (2017) find that the Sino-Tibetan, Kra-Dai, and Hmong-Mien languages are all tonal. Of the Austroasiatic languages, about one-third have register contrasts, a number of which incorporate pitch, and another third have three or more tones (with many also using voice quality concurrently). The Austronesian languages for the most part utilize neither tone nor register, with the exception of a number of languages (almost all Chamic) that utilize pitch in register contrasts. One Chamic language, Tsat, is not strictly in MSEA, but has a 5-tone system (Thurgood, 1999, p. 274). Tonogenesis, the emergence of tone, has occurred in a number of languages. Of relevance to the language families of Southeast Asia is tonogenesis in Chinese sometime between the first millennium BCE and CE (Sagart, 1999, p. 101), and in Kra-Dai, Hmong-Mien, and Viet-Muong around the same time. elucidated through historical and comparative evidence. While some sources attribute tonogenesis, either implicitly or explicitly, in these languages to contact with Chinese due to identical-looking tone systems (Benedict, 1996; Matisoff, 1973, p. 88, Pulleyblank, 1986; Sagart, 1999; Ferlus, 2004, p. 307), there is doubt about the source necessarily being Chinese, the tone systems directly being “borrowed”, and about contact as an explanation in general (Ratliff, 2015a; Brunelle & Kirby, 2015). Ratliff (2015a), however leaves room for the possibility of contact making languages more ‘tone prone’ (p. 261). The existence of tonal languages in generally non-tonal language families but that are in areas with other tonal languages offers circumstantial evidence for this idea, although there are several languages not in close contact with tonal languages that develop tone as well.

Tonogenesis is understood to derive from laryngeal contrasts or, more rarely, from vowel contrasts (see Kingston, 2011 and Ratliff, 2015a for overviews). These contrasts tend to involve the co-occurrence of various cues, one of which is the manipulation of fundamental frequency (f0). Several studies have shown small, but significantly different f0 trajectories of vowels in various phonological contrasts that are not primarily cued by f0.4 Experimental data has also shown generational differences in the usage of f0 cues in what was originally a non-tonal contrast in several languages in the present-day (e.g. Seoul Korean: Kim, 2004; Afrikaans: Coetzee et al., 2018).

The contrast of interest in the current study is that of register. Register is a phonological contrast that employs a constellation of suprasegmental features, including pitch, voice quality, and vowel quality.5 Many languages that have a two-way register distinction may be described as having a modal-breathy, creaky-modal, tense-lax, or stiff-slack distinction, for example, many of which are terms referring to voice quality—the phonetics associated with each of these terms differs in nuanced ways.6 These pairs are often collectively referred to as “high” and “low” register. Table 1, adapted from Brunelle et al. (2020) and added to, summarizes crosslinguistic cues for register from various studies.

Like tonogenesis, registrogenesis, the development of register as a phonological contrast, also has origins in laryngeal contrasts such as voicing and aspiration. The development of a modal-breathy contrast in various Austroasiatic languages, including Kuy, arises from voiced onsets.
Table 1: Crosslinguistic correlates of register contrasts

<table>
<thead>
<tr>
<th>Cue</th>
<th>High Register</th>
<th>Low Register</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open quotient</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Spectral tilt</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Harmonics-to-noise ratio</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>Intensity</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>f0</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>F1</td>
<td>Higher</td>
<td>Lower</td>
</tr>
<tr>
<td>F2\textsuperscript{8}</td>
<td>More peripheral?</td>
<td>More centralized?</td>
</tr>
<tr>
<td>VOT</td>
<td>Shorter</td>
<td>Longer</td>
</tr>
<tr>
<td>Vowel duration</td>
<td>Shorter</td>
<td>Longer</td>
</tr>
</tbody>
</table>

Registrogenesis and tonogenesis come about through an increase in the reliability of “non-primary cues” to signal the phonological contrast, a process that may result in the neutralization of the original cue. As such, the underlying contrast is preserved, but the phonetic cues change, a process termed transphonologization (Haudricourt, 1965; Hyman, 2013): the original segmental contrast becomes a suprasegmental one of register or tone. Transphonologization may begin with enhancement of a secondary cue (in the case of tonogenesis, this would be f0), leading to redundant cueing of the contrast alongside the original cue (voicing or voice quality, for example). Following this redundancy, the original cue may be weakened, eventually leading the originally secondary cue to become primary (Maran, 1973; Hyman, 1976). Alternatively, it is possible that the primary cue first weakens, leading to the potential for merger. In this situation, enhancement of another cue may be a strategy to avoid merger, which may be particularly important if the contrast has a high functional load.

Two large factors that are relevant to sound change are the emergence of structural features and social dynamics that tilt the likelihood of change in a certain direction. As an example of the former, a structural feature that is well-known to be correlated with and potentially a precursor for tonogenesis is monosyllabicity (Michaud, 2012; Kirby & Brunelle, 2017). As syllables reduce, the functional load of each segment increases in turn. If a segmental contrast becomes less salient, there may be a pressure to enhance suprasegmental cues at the risk of merger. A social factor that may serve as an example of the latter is that of language contact, the effects of which are visible at all levels of language, from phonology to syntax to discourse. Structural change due to contact is understood to be a result of bilinguals’ imposition of features from one language they speak upon another one (van Coetsem, 1988; Winford, 2005). If bilingualism is necessary for structural changes to permeate into a language from another, then an understanding of how bilinguals utilize cues differently from monolinguals is vital to explain how contact may bring about change. L2 speakers of a language use have been shown to use cues differently from monolinguals in both production and perception. Effects are not merely directional from the L1 to the L2; L2 knowledge can also affect shift L1 categories in both perception and production at all levels of language, even with short-term or passive exposure.

1.2 The linguistic landscape of Tambon Tum

Both passive exposure and active usage of multiple languages are facets of daily life in Kuy society in Thailand. The Kuy in Thailand primarily live in the southern part of Isan, a common name for Northeast Thailand, and are concentrated in the three provinces of Buriram, Sisaket, and Surin.
all of which border Cambodia. As a region at the crossroads of three modern countries (Thailand, Laos, and Cambodia) and their predecessors, multilingualism has been the norm for centuries. In order to understand the intersection of language usage and sound change, the sociolinguistic background of Kuy and relevant phonology of the four languages used by Kuy speakers will first be presented.

1.2.1 Sociolinguistic background of Kuy

The Kuy in Thailand are historically quadrilingual in Kuy, Khmer, Lao, and Thai. These four languages exist in a hierarchy: in a classification devised by Smalley (1994), Standard Thai [ISO 693-3: tha] sits at the top, being the national language taught in schools. Just under Standard Thai is the regional language, Northeastern Thai [ISO 693-3: tts], encompassing several varieties spoken in Isan that are contiguous with varieties of Lao [ISO 639-3: lao] in Laos and linguistically, all these varieties are grouped together as “Lao”. However, given the political boundary between Thailand and Laos, there are various sociolinguistic differences between the two. The Kuy speak a variety of Southern Lao. Next in the hierarchy is Khmer, spoken along the Thailand-Cambodia border, a marginal regional language with a sizable population of speakers and which is codified as the national language just across the border. The variety spoken by the Kuy population is Northern Khmer [ISO 639-3: kxm]. At the bottom of this hierarchy is Kuy [ISO 639-3: kdt], a marginal language that lacks national status. Kuy is a West Katuic language in the Austroasiatic language family spoken at the border of Thailand, Laos, and Cambodia. It lies in a dialect continuum with varieties known as Kuay (which shares the same ISO 639-3 code) and Nyeu (or Yeu) [ISO 639-3: nyl]. These groups are also known by the exonym Suay. Figure 1 shows the distribution of the Katuic languages from Diffloth (2011, p. 10). The red arrow I have added points to the Kuy variety, spoken in Tambon Tum, in the current study. While Diffloth does not mark country borders, the reader may use Ubon, Pak Se, and Preah Vihear as landmarks in Thailand, Laos, and Cambodia, respectively.

For brevity, and clarity, the four languages in this study will be generally be referred to as Kuy, Khmer, Lao, and Thai, while admitting that these terms homogenize and oversimplify linguistic and social distinctions. Generally, an ethnolinguistic group can speak most, if not all, the languages in the area that are higher than their own language on Smalley’s hierarchy, but not ones that are lower; as such, many Kuy speakers have command of all four languages in the area. The Kuy have historically assimilated to neighboring Khmer and Lao groups, leading to shrinking of the Kuy-speaking region in recent times, despite growth of the population (Seidenfaden, 1952; Yantreesingh, 1980, p. 3; Smalley, 1988, p. 396). Estimates of the total number of speakers in this continuum across Thailand, Laos, and Cambodia is between 348,000 to 473,000, with 275,000 to 400,000 living in Thailand (Smalley, 1994; Chazée, 1999; Lefebvre, 2000; Premsrirat, 2006). At the time of Premsrirat (2006), the Kuy had an affection for their language, but the use of Lao was seen as more prestigious and was generally the language of choice outside the home.

Demographic data looking at the progression of language shift may be seen in Tables 2, 3, and 4. These data are from 117 speakers interviewed in 2018 and 2019. The results are split by generation, with 45 years old considered the cutoff. Table 2 breaks down speakers’ ability in Khmer and Lao and shows a generational shift in the degree of multilingualism: while the older generation is fully tri- or quadrilingual, there are much fewer quadrilingual speakers in the younger generation and there are even 4 individuals who are only bilingual in Kuy and Thai. Table 3 shows self-assessment of speaking ability in each language as compared to Kuy. The most notable generational shift is the switch from the majority of speakers rating their Thai ability as less than their Kuy ability (31/59) to the majority of speakers rating their ability equally in the two
languages (31/58). Table 4 shows self-assessment of frequency of language usage as compared to Kuy. The numbers here show similar pattern to speaking ability: the large majority of older speakers use Kuy more than Thai (45/58). While the number of younger speakers who use Kuy more than Thai is still the majority (29/58), there is a large increase in speakers using Thai more than Kuy (from 3/58 to 26/58). There also appears to be a slight increase in speakers using Lao more than Kuy (from 1/58 to 7/58). While past literature has described Kuy as losing ground to Khmer or Lao, the current numbers show that in recent times, it is primarily Thai that speakers are shifting to.

Table 2: Lao and Khmer speaking ability (based on author’s survey data)

<table>
<thead>
<tr>
<th>Age</th>
<th>Neither</th>
<th>Khmer only</th>
<th>Lao only</th>
<th>Both</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;45</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>55</td>
<td>59</td>
</tr>
<tr>
<td>≤45</td>
<td>4</td>
<td>2</td>
<td>14</td>
<td>38</td>
<td>58</td>
</tr>
</tbody>
</table>

Table 3: Speaking ability compared to Kuy (based on author’s survey data)

<table>
<thead>
<tr>
<th>Age</th>
<th>Thai less</th>
<th>Thai same</th>
<th>Thai more</th>
<th>Khmer less</th>
<th>Khmer same</th>
<th>Khmer more</th>
<th>Lao less</th>
<th>Lao same</th>
<th>Lao more</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;45</td>
<td>31</td>
<td>25</td>
<td>3</td>
<td>49</td>
<td>9</td>
<td>1</td>
<td>44</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>≤45</td>
<td>14</td>
<td>31</td>
<td>13</td>
<td>53</td>
<td>4</td>
<td>1</td>
<td>45</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>
Ban Khi Nak School timeline (EMIS 2013, p.c. Sidawun Chaiyapha)

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1939</td>
<td>School established (up to 4th grade)</td>
</tr>
<tr>
<td>1972</td>
<td>Expanded to 6th grade</td>
</tr>
<tr>
<td>1998</td>
<td>Expanded to 9th grade</td>
</tr>
</tbody>
</table>

Table 4: Language usage frequency compared to Kuy (based on author’s survey data)

<table>
<thead>
<tr>
<th>Age</th>
<th>Thai less</th>
<th>same</th>
<th>more</th>
<th>Khmer less</th>
<th>same</th>
<th>more</th>
<th>Lao less</th>
<th>same</th>
<th>more</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;40</td>
<td>45</td>
<td>10</td>
<td>3</td>
<td>54</td>
<td>4</td>
<td>0</td>
<td>54</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>≤40</td>
<td>29</td>
<td>3</td>
<td>26</td>
<td>55</td>
<td>2</td>
<td>1</td>
<td>49</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

The increasing shift to Thai in recent times is not unique to the Kuy, but is prevalent throughout Thailand, as a result of continued efforts to centralize the country. Three salient changes that have occurred over the past several decades are longer schooling, an increase in interregional travel, and greater consumption of central Thai media. Mandatory primary schooling was introduced in 1921 and schools have played a central role in the development of a national and unified Thai identity through emphasis of loyalty to three symbols—the King, the nation, and Buddhism (Smalley, 1994, p. 323, McCargo & Hongladarom, 2004)—and through the enforcement of Standard Thai in schools beginning in the reign of King Vajiravudh (1910–1924) (Suraratdecha, 2014, p. 240). As a result, younger generations are becoming increasingly fluent in Standard Thai and mobile, moving to larger cities for job opportunities. Migration of Isan people to larger cities as “cheap, unskilled labor” grew following World War II, although gender dynamics have changed over time: in the 1960s, rural migrants were mostly males without spouses or children, but by the 1990s, consisted of all ages and genders (Hesse-Swain, 2011, p. 44).

Ban Khi Nak School is the school that was attended by almost all participants in this sample. Teachers in the past were primarily hired from the local area, so students could use Kuy to varying extents in school with the expectation that teachers would at least understand. However, this expectation no longer holds, as teachers are now mainly recruited from other parts of the country and may not even be aware of the existence of the Kuy as a separate ethnolinguistic group. Table 5 shows the timeline of the establishment and expansion of Ban Khi Nak School.

As schooling was much more limited in the past, many of the older speakers in this sample did not complete school past fourth or sixth grade, as the school was not expanded to ninth grade until just over twenty years ago. Most of the younger population in this study also attended high school nearby and continued to college. Those who continue to college are immersed in Thai by virtue of being in school longer and by living in non-Kuy speaking areas. Lao also serves as a regional lingua franca for those who come from various parts of Isan. Following college, much of the younger population proceed to work in other regions of Thailand, continuing the aforementioned trend of Isan migrants into larger cities, where they use Thai and/or Lao. This continuing trend is facilitated by the continued rapid improvement in transportation infrastructure in Thailand. While spending time in other parts of Thailand was not uncommon in the older generation, it has become much more mainstream. Table 6 shows the time participants have spent away from home, divided by age. The difference by generation is stark—while the majority of older people have spent no time away from home, this is the case for only seven people in the younger generation.
### Table 6: Time spent away (100 speakers, data from current study)

<table>
<thead>
<tr>
<th>Age</th>
<th>Has not spent time away</th>
<th>Has spent time away</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>M</td>
</tr>
<tr>
<td>&gt; 45</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>≤ 45</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

### Table 7: Consonants of Kuy (adapted from Phimjun (2004, p. 24–25))

<table>
<thead>
<tr>
<th>Labial</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p pʰ b</td>
<td>t tʰ d</td>
<td>c cʰ</td>
<td>k kʰ</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>ɲ</td>
<td>ŋ</td>
</tr>
<tr>
<td>Fricative</td>
<td>s</td>
<td></td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>Trill</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>w¹⁶</td>
<td>j</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The shift in the gender balance of migrants pointed out by Hesse-Swain (2011, p. 44) is mirrored in this data as well: in the older generation, 46.67% of females had spent any time away from home, as opposed to 65.52% of males, while in the younger generation, this gap closes with the proportion being 86.21% for females and 89.66% for males.

The demographic information above shows a very clear generational shift in the Kuy population in this study: the younger generation is spending more time in school and in other parts of Thailand. As a result, they use much more Thai and many consider themselves to be equally bilingual in Kuy and Thai, a finding mirrored in Siebenhütter (2020). The language dynamics are changing in Kuy society due to both the average life trajectory of a member of the Kuy community and to the distribution of languages that are now heard in the village, due to the encroachment of Thai into more linguistic contexts. Many younger participants report using Thai, rather than Kuy, in the home with their children, with a common reason being to expose their children to Standard Thai early for the purposes of succeeding in school. Many younger parents report that their children can understand, but not speak, Kuy. Due to these shifts in Kuy usage, it is currently classified by Ethnologue as 6b: Threatened (Simons & Fennig, 2017) and as “severely endangered” by UNESCO (Moseley, 2010).

#### 1.2.2 Phonology of Kuy and surrounding languages

The current study focuses on Kuy as spoken in Tambon Tum of Amphoe Prang Ku in the Sisaket province of Northeastern Thailand. Discussion of phonology in this section will be limited to facts relevant to the current study, but the reader may refer to Sriwises (1978), Yantreesingh (1980), Suwannaraj (1990), Sangmeen (1992), Sukgasame (2003), Phimjun (2004), or Gehrmann (2016) for fuller descriptions on different Kuy, Kuay, and Nyeu varieties. The consonants and vowels of Kuy as laid out by Phimjun (2004) may be found in Tables 7 and 8. Phimjun’s inventory is based off data elicited from Kuy as spoken in Tambon Ku, which borders Tambon Tum to the south. This inventory matches what I observed in the Kuy of Tambon Tum. Kuy has a two-way register contrast between modal and breathy voice, a common feature of Austroasiatic languages (Jenny & Sidwell, 2014, p. 53). Phimjun does not discuss whether all vowels are represented in both registers; however, Gehrmann (2016) reviews multiple sources,
confirming that different Kuy varieties have different restrictions, due to registrogenesis and subsequent vowel changes occurring after Proto-Kuay (pp. 36–48). Kuy in Tambon Tum displays gaps as well, but I have not yet determined all of them. Voiceless unaspirated and aspirated stops are neutralized before breathy vowels, a common feature of Katuic (Huffman, 1976; Diffloth, 1982; Gehrmann & Kirby, 2019). Sriwises (1978, p. vii) and Sukgasame (1993, p. 249) describe modal voice as having a higher pitch overall than breathy voice. Alongside f0 differences, L. Thongkum (1989) also finds higher amplitude in modal voice but conflicting patterns for differences in the first formant (F1) and duration.

Kuy monomorphemic words are maximally of the shape C(V/N).CRVC, in which C stands for a consonant, N a syllabic nasal, and R a liquid. The first syllable is limited in three ways: (1) The onset must be simplex, (2) the nucleus must be either a minimal vowel that ranges between nothing and a schwa or a syllabic nasal homorganic with the onset of the second syllable (to be represented as N), and (3) there is no coda. This word shape with an unstressed and restricted first syllable (often called a “minor syllable”) is common in Austroasiatic and is termed sesquisyllabic, meaning “syllable and a half” (Henderson, 1952; Matisoff, 1973). One example of a word with the maximal syllable structure is /cN̩tray/ “diligent”. The presence of the syllabic nasal shows age-graded variation in Tambon Tum, such that it is often dropped by younger speakers but rarely by older ones. Reduction and disappearance of the syllabic nasal has implications for tonogenesis, to be discussed in 4.4.

The phonological details for Thai, Lao, and Khmer will be restricted to facts relevant to this study. Standard Thai has five tones: mid, low, high-falling, high-rising, and low-rising (Abramson, 1962; Tingsabadh & Deeprasert, 1997; Iwasaki & Ingkaphirom, 2005). Closed syllables with a stop coda may only take one of two tones: if the vowel is short, the tone may be low or high; if it is long, the tone may be low or high-falling. Depending on the variety, Southern Lao may have five or six tones—the Sisaket variety has five tones in open syllables: low-rising, high-falling, high, glottalized low, and glottalized mid-falling. Closed syllables with a stop coda may be mid-rising or high if the vowel is short and glottalized low or glottalized mid-falling if it is long (Brown, 1965; Hoonchamlong, 1984, Sipipattanakun, 2014, p. 109). Northern Khmer has neither tone nor register. While most native words in both Thai and Lao are monosyllabic, many sesquisyllabic loanwords have entered the language through Khmer and polysyllabic ones through Sanskrit, Pali, English, and French (in the case of Lao). Khmer, on the other hand, shares the typical Austroasiatic syllabic structure with Kuy, having many sesquisyllabic and monosyllabic words. Like Thai and Lao, it also has many polysyllabic words from Sanskrit, Pali, English, and French.
1.3 Usage of f0 in Phonological Contrast

As mentioned earlier, modal voice tends to have higher f0 than breathy voice in Kuy. The continual efforts to centralize Thailand increase exposure to and usage of Thai. If these dynamics sufficiently increase the intensity of contact, we may expect cue weights to flow in the direction of f0. There are a number of contemporary studies on language contact effects on f0 usage in the context of the Mainland Southeast Asia linguistic area. Brunelle (2009) explores cues in the register contrast among three dialects of Cham, demonstrating that Eastern Cham speakers, who are highly bilingual in tonal Vietnamese, show the greatest pitch differences. In exploring the realization of Lao tones by speakers with different language backgrounds is Isan, Pratankiet (2001) finds that Khmer and Kuy speakers show citation form differences from bilingual Lao-Thai speakers20, while Sipipattanakun (2014) shows that they have narrower f0 pitch ranges than those who are only bilingual in Lao and Thai. Both linguists attribute these differences to the lack of tone in Kuy and Khmer, but point out as well that differences in some tone realizations may be attributable to influence from Standard Thai.

The effect of Kuy bilingualism on the Thai tone contrast shows substrate effects on a superstrate language. Potential evidence for the reverse comes from a number of studies on the f0 contrast in Kuy. An apparent time study on three Ku(a)y varieties in Thailand and one in Laos by Sukgasame (2003) reveals that the Thai varieties are giving way to a pitch distinction and the Lao ones to a vowel quality distinction in younger speakers. In a follow-up study on two Thai varieties (one of which overlapped with Sukgasame, 2003), Sukkasame (2004) shows similarities between the emergent pitch patterns in these varieties and the tone patterns in the neighboring Lao varieties. A production and perception study by Abramson et al. (2004) finds that voice quality is a weak cue in production and perception for some Kuy speakers. They suggest that Kuy is shifting from a register distinction to one with accentual salience and that this may be due to influence from Thai or Lao.

Thus, the literature describes varieties of Kuy in which the register contrast may be shifting to one with greater usage of pitch or giving way to the incipience of tonogenesis. Expanding upon Lau-Preechathammarach (in press), which demonstrated greater usage of f0 cues for Kuy speakers who use Thai or Lao more than for those who use it less at the group level in both perception and production, this study delves into individual differences in the acoustic correlates of the Kuy register distinction and examines whether the interaction with language usage and other social factors found at the group level is also found at the individual level. Furthermore, the interaction between f0 and voice quality at the individual level is also investigated in order to understand whether there is a tradeoff between the two cues. The goal of this study is to provide a close analysis of how we might understand the macro-effects of language contact on linguistic structure in language used by a community through the micro-effects of bilingualism on the realization of a phonological contrast at the individual level.

2 Hypothesis and Methods

As Thai and Lao are tonal languages that primarily use f0 for phonological contrast, participants who use Thai/Lao more are hypothesized to also use f0 cues more and other voice quality cues less in the register contrast, as compared to participants who do not use Thai/Lao as much. A trading relation is also expected between f0 and other voice quality cues, such that speakers who weigh f0 more heavily are expected to weigh other voice quality cues less heavily.

In order to test these hypotheses, a production study was carried out in which participants
embedded modal and breathy minimal pairs in a carrier sentence. Following the task, speakers were interviewed to gather demographic and sociolinguistic information. Acoustic measures were taken over the course of the target vowels for analysis. Data was transformed by using a principal component analysis to highlight the most meaningful dimensions of variation. Linear regression models were then fitted to estimate the general effects of sociolinguistic factors on trajectories of the acoustic measures. Finally, a linear discriminant analysis was run to determine the acoustic cue weights for each individual and correlations between cue weights and sociolinguistic factors were analyzed to understand within-speaker patterns.

2.1 Participants

In Tambon Tum, there are three villages with a sizable Kuy population: Ban Khi Nak (Kuy: /tʰɔʔ kʰna:k/), Ban Rong Ra (Kuy: /tʰɔʔ araʔ/), Ban Khi Nak Noi (Kuy: /tʰɔʔ kɛ:t/). According to Kuy speakers in Tambon Tum, these three villages speak the same variety of Kuy. 75 participants were recruited from these three villages with the help of Thongwilai Intanai, a Kuy speaker from Ban Khi Nak. Participants were explicitly balanced for age and gender, comprising four decades (twenties, thirties, fifties, and sixties) and two genders (female and male). At least eight speakers were sought for each age-gender combination, but given time constraints, extra speakers were also recruited opportunistically, such that some subgroups are overrepresented. Ultimately, nine speakers were excluded from analysis due to failing to complete the experiment (one), extreme difficulty with the task (two), recording issues (four), and producing fewer than ten analyzable (see §2.5 for determination of analyzability) unique words (two), leaving 66 participants (40 from Ban Khi Nak, 14 from Ban Rong Ra, and 12 from Ban Khi Nak Noi). These participants’ ages (divided by decade) and genders are provided in Table 9.

Table 9: Participants by age and gender

<table>
<thead>
<tr>
<th></th>
<th>20s</th>
<th>30s</th>
<th>50s</th>
<th>60s</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>M</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>9</td>
</tr>
</tbody>
</table>

A sociolinguistic questionnaire, a translated version of which may be found in Appendix D, designed to capture factors related to language ability, language usage frequency, time spent away from home, and ethnolinguistic affiliation, was included in the study. Apart from age and gender, seven other self-reported sociolinguistic factors, described in Table 10, were focused on for this study. The languages and ethnolinguistic groups asked about in the questionnaire were Kuy, Thai, Lao, and Khmer.

All speakers are bilingual in Kuy and Thai and most have at least some knowledge of Lao and Khmer. Only one participant reports not understanding Lao, while six report not understanding Khmer. In terms of speaking ability, three participants report not being able to speak Lao, while 16 report not being able to speak Khmer. All participants are able to read and write Thai, although some older participants have more difficulty doing so. The distribution of time spent away from home (measured in years) may be seen in Figure 2. While 20 participants (almost one-third) report never having lived outside of Tambon Tum, there is a fair spread in the time spent away for the remaining 46 participants. The mean and median number of years spent away from home are 5.7 and 4, respectively, while the minimum (non-zero value) and maximum numbers are 0.5 and 47, respectively. There is a considerable amount of variation in language usage and ability
Table 10: Sociolinguistic Variables

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years Away</td>
<td>Years spent living in another (non-Kuy speaking) area</td>
</tr>
<tr>
<td>Overall Frequency</td>
<td>Overall frequency of using language (0% = never, 100% = all the time)</td>
</tr>
<tr>
<td>Friend Frequency</td>
<td>Frequency of using language with friends (0% = never, 100% = all the time)</td>
</tr>
<tr>
<td>Family Frequency</td>
<td>Frequency of using language with family (0% = never, 100% = all the time)</td>
</tr>
<tr>
<td>Comprehension Level</td>
<td>Ability to understand language (0 = understands nothing, 4 = understands fully)</td>
</tr>
<tr>
<td>Speaking Ability</td>
<td>Ability to speak language (0 = cannot speak, 4 = speaks fluently)</td>
</tr>
<tr>
<td>Strength of Identification</td>
<td>Identification with ethnolinguistic group (0 = does not identify, 3 = identifies strongly)</td>
</tr>
</tbody>
</table>

Figure 2: Histograms of time spent away by participants

and in time spent away from home. It is the relationship between this variation and manifestation of the acoustic correlates of register that the current study seeks to probe.

2.2 Stimuli

The stimuli for the experiment consisted of 58 unique words, consisting of 31 target words and 27 distractor tokens (see Appendix A). The target words comprised 12 modal unaspirated vs. breathy pairs, 2 modal aspirated vs. breathy pairs, and 1 modal unaspirated vs. modal aspirated vs. breathy triplet. Of the target words, 4 potentially have syllabic nasals. Words were balanced only for voice quality, but not for segments. The remaining distractor tokens were chosen to observe phenomena for future research—of these, two variables of sociolinguistic interest are /tr/ clusters and final /r/, which show age-graded changes like the dropping of syllabic nasals. Historical /tr/ is realized as [kr] in a fair number of mostly younger participants in the study, while three of the words in this study with final /r/ (/piːr/ ‘flower’, /piːr/ ‘to wind’, /sN̩kiːr/ ‘to be sensitive’) are realized with final [l] in almost all speakers. These two variables show similar variation in other Ku(a)y varieties (Sukkasem, 2005, p. 50).
2.3 Procedure

The task in the current production study involved embedding the stimuli in a carrier sentence. The sentence was presented in Thai primarily because Kuy orthography is recently developed and few speakers are familiar with it, but secondarily to avoid reading pronunciation. The example sentence presented to participants is in (1). In order to facilitate naturalness, participants were asked to translate this sentence into Kuy and to use that frame for each sentence, only replacing the word “water” with the stimulus. One common translation of the sentence is in (2). Differences mainly lay in presence or absence of the complementizer /paj/ ‘say’ and uses of synonyms, such as /kʰam/ for /pna:j/ or /pa:j/ for /waw/. The stimulus retained prominence regardless of the sentence used. The experiment was presented to participants on a Google Nexus 10 tablet and carried out in a quiet room in the temple of Wat Nakharin in Ban Khi Nak. Participants were recorded using a C544-L head-worn condenser microphone connected to an H4n Zoom recorder. Pictures were included with each sentence to aid in elicitation of the intended word. Figure 3 shows the screen that participants saw for the example sentence with “water”, while Figure 4 shows what participants saw for the trial sentence with “egg”.

(1) Example Thai sentence

ฉัน ชื่อ คำ ว่า น้ำ ให้ เขา ฟัง

'Isaytheword “water” for them to hear.’

(2) Example Kuy translation (in Mahidol orthography)

ไส ว่า ประโยค ไป ผู้ คุณ เขา จะได้

'Isaytheword “water” for them to hear.’

In the training round, the participant first settled on an appropriate translation for the sentence.
We then went through each stimulus once as a practice round for familiarization. The participant then completed five rounds alone (generally taking between 15 and 30 minutes), with an optional break after the third round. Stimuli were randomized in both the practice and trial rounds. Participants were told that they could skip a word if they could not remember it. As there were 31 target words and five rounds, 155 targets words in total could theoretically be produced.

### 2.4 Preprocessing of sociolinguistic information

Principal component analysis (PCA) is a technique used to reduce a number of variables to a small number of orthogonal dimensions that capture variance between the variables and is particularly useful when variables are expected to be highly correlated with each other. As much of the sociolinguistic information was expected to be correlated to language usage, a PCA was carried out using the FactoMineR package (Lê et al., 2008) in R (R Core Team, 2018) to reduce the data to a tractable number of variables for analysis. Many of the variables (see Table 10) involved frequency of usage or level of ability/identification, all with respect to the 4 languages/ethnic groups in the area. As such, this data was first consolidated by calculating differences between answers with respect to Kuy and to the tonal languages, Thai/Lao, rather than taking the reported values at face value for two reasons. First, participants differed in how “humble” they were with their answers. For example, some participants rank their ability as low in even their most comfortable language, while others are more “confident”. Second, the goal was to yield a proxy measure that would capture the relative level of Kuy usage/ability/identification as compared to Thai/Lao.

As can be seen in Figure 2, there is a heavy right skew for the number of years away people have spent away from home (mean: 5.7, median: 4, skew: 3.12\(^2\)). Due to this skew, the number of years spent away was square rooted. The distribution of \(\sqrt{\text{Years Away}}\) can be seen in Figure 5 (mean: 1.84, median: 2, skew: 0.57).

![Figure 5: Histogram of square root of time spent away by participants](image)

The transformed sociolinguistic variables that were input into the PCA are explained in Table 11. Values were scaled by dividing z-values by two, following Gelman (2008), who proposes this method for statistical modeling to allow for direct comparison with binary predictors.
Table 11: Sociolinguistic Variables for PCA

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Participant’s age in years.</td>
</tr>
<tr>
<td>√Years Away</td>
<td>Square root of years spent living in another (non-Kuy speaking) area.</td>
</tr>
<tr>
<td>Understand</td>
<td>Ability to understand Kuy (coded from 0-4) minus ability to understand Thai or Lao (coded from 0-4; the higher ranking between the two languages is chosen).</td>
</tr>
<tr>
<td>Speak</td>
<td>Ability to speak Kuy (coded from 0-4) minus ability to speak Thai or Lao (coded from 0-4; the higher ranking between the two languages is chosen).</td>
</tr>
<tr>
<td>Overall Freq</td>
<td>Overall frequency of using Kuy (coded from 0-100) minus overall frequency of using Thai or Lao (coded from 0-100; the higher number between the two languages is chosen).</td>
</tr>
<tr>
<td>Family Freq</td>
<td>Frequency of using Kuy with family (coded from 0-100) minus frequency of using Thai or Lao with family (coded from 0-100; the higher number between the two languages is chosen).</td>
</tr>
<tr>
<td>Friend Freq</td>
<td>Frequency of using Kuy with friends (coded from 0-100) minus frequency of using Thai or Lao with friends (coded from 0-100; the higher number between the two languages is chosen).</td>
</tr>
<tr>
<td>ID</td>
<td>Self-rating of Kuy identity (coded from 0-3) minus self-rating of Thai or Lao identity (coded from 0-3; the higher number between the two languages is chosen).</td>
</tr>
</tbody>
</table>

2.5 Acoustic Analysis

Following forced alignment on the production results using the Montreal Forced Aligner (McAuliffe et al., 2017), target vowel boundaries were realigned and, for stops, the voice onset time (VOT) boundaries were marked by myself and twelve undergraduate research assistants. For each token, eight acoustic measures were taken over the course of the target vowel using VoiceSauce, a software for calculating voice measures (Shue et al., 2011). These measures are f0, F1, CPP, and five harmonic measures ($H_1^\star, H_1^\star-H_2^\star, H_1^\star-A_1^\star, H_1^\star-A_2^\star, H_1^\star-A_3^\star$). CPP, cepstral peak prominence, is a proxy for harmonics-to-noise ratio (HNR). Breather voice qualities have lower HNR and CPP values (Hillenbrand et al., 1994). $H_n$ refers to the amplitude of the $n^{th}$ harmonic, while $A_n$ refers to the amplitude of the loudest harmonic in the $n^{th}$ formant. $H_1 - H_2$ correlates with open quotient, the ratio of the glottal cycle for which the vocal folds are open, while the $H_1 - A_n$ values are measures of spectral tilt, the rate of the loss of energy as the frequency of harmonics increase. Higher open quotients and steeper spectral tilt are associated with a breathier voice quality (Holmberg et al., 1995; Hanson, 1995, 1997; Hanson & Chuang, 1999; Henrich et al., 2001; Gobl & Ní Chasáide, 2013). The asterisks following the spectral measures indicate correction for formant frequencies and bandwidths and in order to account for age and gender differences between speakers (Iseli et al., 2007). All of the spectral tilt measures were ultimately combined into one, to be called $H_1^\star(-A_n^\star)$, through a PCA on all the acoustic variables except for f0 and F1 (see §3.2.2), leaving five measures for analysis: f0, F1, CPP, $H_1^\star-H_2^\star$, and $H_1^\star(-A_n^\star)$.

Files were marked for whether the uttered word was the intended target and, for the four words with potential syllabic nasals, whether the nasal was existent. Minimal pairs were included in the analysis only if the speaker produced at least two tokens of each member of the pair and only if the tokens for each member of the pair matched in presence or absence of the syllabic
nasal. For most pairs, this meant that the word with the potential syllabic nasal had to lack it (ex. /N̩cʰuːn/ ‘to hide’ and /cʰuːn/ ‘to send’), but in the case of /N̩kɛːŋ/ ‘waist’ vs. /(N̩)kɛːŋ/ ‘side’, the pair was included only if both words had at least two tokens matching in presence or absence of the nasal. Following this procedure, 5125 files were available for analysis.

Because voice tracking algorithms are sensitive to individual differences, f0 and the first three formants were calculated with speaker-specific parameters. f0 was tracked with the STRAIGHT algorithm (Kawahara et al., 1998) through VoiceSauce. Pitch halving and pitch doubling errors were identified and the pitch floor and ceiling were adjusted accordingly. Formants were measured with Praat (Boersma, 2001), also through VoiceSauce. For participants whose f0 values averaged below 150 Hz, the formant ceiling was set at 5500 Hz and the number of formants to be detected was set at 5.5, while for those whose f0 values averaged above 150 Hz, these values were set at 5000 Hz and 4.5, respectively, following heuristics laid out by Skarnitzl et al. (2015). After inspection, the ceiling was shifted as necessary to minimize errors. These f0 and formant values were the basis for VoiceSauce’s calculation of the five harmonic measures and CPP. All measurements were taken at every millisecond with a sliding window of 25 ms.

Three more steps were implemented to handle errors. First, a moving median filter with a size of 15 ms was applied to smooth out sudden tracking jumps. Second, zero values (0.1% of the dataset) were removed. Finally, z-scores were calculated, using the means and standard deviations for each combination of speaker × vowel quality × voice quality, and values greater than 3 standard deviations away from the mean (5.2% of the remaining dataset) were removed, minimizing tracking jumps that persisted for longer than the median filter window. Because individuals speak at different rates and vowels vary in their length, time was normalized by binning measurements for each file into twenty time intervals and calculating the mean value at each interval. Ultimately, 5.29% of the dataset was removed. Voice quality measures vary largely across individuals. For this reason, the values were scaled by speaker, with Gelman’s 2008 standardization procedure (see §2.4), for comparability and statistical modeling. For F1, the values were additionally scaled by vowel height. f0 and F1 Hertz values were also converted into semitones to better approximate auditory distance (Nolan, 2003).

In order to explore group differences and the overall relationship between sociolinguistic factors and the voice quality measures, linear mixed effects regression models were fitted for the scaled values of each voice quality measure with the lme4 (Bates et al., 2015) package in R. The model looked at the interaction of Timepoints (modeled with B-splines using three knots to capture smooth curves with knots at arbitrary timepoints (Curry & Schoenberg, 1966)), Register, Gender, Tonal Language Experience, and Time Away (see §3.2.1 for an explanation of the last two predictors). A random intercept was included for Word, but random slopes were not included as they led to overfitting.

Individual differences were then explored by quantifying cue weights for each voice quality measure through the use of Linear Discriminant Analysis (LDA), a technique that involves training a classifier to categorize stimuli into classes given a set of information and labels for that information. LDA, carried out through the scikit-learn library in Python 3.7 (Van Rossum & Drake, 2009) is the method of choice as it has been shown to be a robust technique for approximating cue weights and therefore is useful for analyzing individual differences in both production and perception studies (Idemaru et al., 2012; Schertz et al., 2015; Schertz & Clare, 2020). The training and test process involved k-fold cross-validation, in which the dataset is split into k folds, or equal subsets, whose size is $\frac{n}{k}$ rounded up or down to the nearest integer as necessary, where $n$ is the size of the dataset. Each of these subsets comprises a test set, while the subset of tokens that excludes the test subset comprises the training set. The classifier is trained on each training set and then applied to the test set. The resulting classification for the test subset is compared to
the actual classification to yield an accuracy score. In this study, a classifier was trained, for each individual, to determine whether each token was modal or breathy, given the acoustic measures for the token. As in the linear regression analyses, the scaled values for all the measures are used. Semitones are used for f0 and F1. A combined accuracy for all the non-f0 measures \( (H_1^* - H_2^*, H_1^*(-A_n^*), CPP, F1) \), henceforth called Voice Quality (VQ), was also obtained by providing the LDA classifier information from these measures. 10-fold cross-validation was used and the mean of the 10 results was calculated as a proxy for the cue weight of each measure. Pearson’s \( r \) was calculated between these cue weights and Tonal Language Experience to test whether there was a correlation between language usage and production cue weights. Correlations between f0 and voice quality cue weights were also calculated in order to test whether there is a trading relation between f0 and voice quality.

3 Results

3.1 Summary Statistics

3.1.1 Voice Onset Time

Of the 5125 files, 4542 contained a target word with a stop onset. The means and standard deviations of the voice onset time (VOT) of these stops are summarized in Table 12 and visualized in Figure 6. While the VOT of breathy vowels sits between those of modal vowels following unaspirated stops and modal vowels following aspirated stops, it is much closer to the former (+9.86 ms) than the latter (-39.58 ms).

![Figure 6: VOT means by stop type and register](image)

**Table 12: Mean VOT durations (ms)**

<table>
<thead>
<tr>
<th>Stop type &amp; Register</th>
<th>Mean</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaspirated modal</td>
<td>16.89</td>
<td>8.68</td>
<td>1938</td>
</tr>
<tr>
<td>Breathy</td>
<td>26.75</td>
<td>15.86</td>
<td>2262</td>
</tr>
<tr>
<td>Aspirated modal</td>
<td>66.33</td>
<td>31.13</td>
<td>342</td>
</tr>
</tbody>
</table>

3.1.2 Duration and f0 by vowel

Table 13 and Figure 7 break down and visualize the means and standard deviations of vowel durations from the data by vowel and voice quality. Scaled values are used because of interspeaker differences. Short breathy vowels are longer than modal ones, matching the generalization that breathy vowels tend to be longer than modal ones, but long vowels do not have a consistent pattern. This inconsistency matches the findings of L. Thongkum (1989) for Kuy, who suspects...
that the durational difference between registers may not occur robustly in languages with a vowel length contrast.

![Figure 7: Duration by vowel quality](image)

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Modal Mean</th>
<th>Modal SD</th>
<th>Modal n</th>
<th>Breathy Mean</th>
<th>Breathy SD</th>
<th>Breathy n</th>
</tr>
</thead>
<tbody>
<tr>
<td>i:</td>
<td>263.42</td>
<td>105.84</td>
<td>447</td>
<td>277.71</td>
<td>111.06</td>
<td>429</td>
</tr>
<tr>
<td>u:</td>
<td>235.94</td>
<td>101.8</td>
<td>718</td>
<td>211.51</td>
<td>114.43</td>
<td>733</td>
</tr>
<tr>
<td>e:</td>
<td>301.62</td>
<td>107.05</td>
<td>228</td>
<td>305.17</td>
<td>98.42</td>
<td>211</td>
</tr>
<tr>
<td>o:</td>
<td>201.23</td>
<td>64.6</td>
<td>440</td>
<td>214.74</td>
<td>67.83</td>
<td>430</td>
</tr>
<tr>
<td>ɛ:</td>
<td>202.16</td>
<td>65.83</td>
<td>61</td>
<td>200.85</td>
<td>62.78</td>
<td>55</td>
</tr>
<tr>
<td>a</td>
<td>116.16</td>
<td>36.29</td>
<td>639</td>
<td>140.62</td>
<td>44.9</td>
<td>648</td>
</tr>
<tr>
<td>ɑ</td>
<td>89.26</td>
<td>20.18</td>
<td>48</td>
<td>106.03</td>
<td>22.61</td>
<td>38</td>
</tr>
<tr>
<td>Overall</td>
<td>207.4</td>
<td>104.17</td>
<td>2581</td>
<td>211.12</td>
<td>104.61</td>
<td>2544</td>
</tr>
</tbody>
</table>

Figure 8 shows mean f0 trajectories and 95% confidence intervals for each vowel, split by generation. The older generation consists of speakers between 50 and 70, while the younger one consists of speakers between 20 and 40. Scaled values are used to normalize for interspeaker differences. Confidence intervals for /ɛ:/ and /ɑ/ are large because they are represented by few tokens. While some register languages show f0 differences in only some vowels\(^{30}\), it is clear that modal voice is higher overall than breathy voice in Kuy, regardless of vowel quality. The other notable pattern is that differences in the younger generation are generally larger and persist over more of the vowel.

![Figure 8: f0 trajectories by vowel quality](image)

### 3.1.3 f0 trajectories by environment

f0 trajectories are observed with respect to segments in the word and generational differences. Figure 9 looks at onset effects on scaled f0 trajectories, taking into account register. f0 following aspirated stops is higher than after unaspirated stops, and in turn higher than after breathy stops.
f0 following sonorants is higher overall. Contour shapes do not differ much by generation, but starting and ending values do, with modal stop trajectories being slightly higher and sonorant trajectories being slightly lower in the younger generation. Note that aspirated stops are represented only by three words and sonorants only by four words, however, so the aspirated and sonorant f0 trajectories may not be representative.

Figure 10 looks at coda effects on f0 values, taking into account voice quality, in order to compare with the findings by Sukkasame (2004) of incipient tonal contours in two Kuay varieties, summarized in Table 14. Sukgasame finds six tonal patterns that are conditioned by register and the rhyme. -N stands for a nasal coda, while -T stands for a stop coda, which may be preceded by a short (V) or long (V:) vowel. The tones described by Sukgasame are accompanied by Chao numerals to indicate the level and shape of the contours.

Table 14: Tonal contours in Kuay described by Sukkasame (2004, p. 693–695) (adapted)

<table>
<thead>
<tr>
<th>Register</th>
<th>Rhyme or coda type</th>
<th>-∅, -N</th>
<th>-h, -VT</th>
<th>-ʔ, -V:T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modal</td>
<td>high falling</td>
<td>[451]</td>
<td>[44]</td>
<td>[45]</td>
</tr>
<tr>
<td>Breathy</td>
<td>low rising falling</td>
<td>[121]</td>
<td>[254]</td>
<td>[22]</td>
</tr>
</tbody>
</table>
For younger speakers, the tones generally match Sukgasame’s findings in terms of contour (with the exception of breathy /-ʔ/ and /-V:T/, which show a steep rise), although the levels differ. Modal /-h/ and /-VT/ are mid rather than high, while breathy /-ʔ/ and /-V:T/ show a low rising tone that ends high. The shape of f0 for breathy vowels generally matches between older and younger speakers, but /-VT/ shows a contour more similar to /-ʔ/ than to /-h/ and /-V:T/ does not quite match /-ʔ/, beginning much lower instead.

What is particularly striking is that the younger generation has virtually fully overlapping contours for /-h/ and /-VT/ as well as for /-ʔ/ and /-V:T/, showing six tone contours as described in the Kuay varieties observed by Sukgasame.31 The consolidation of these f0 trajectories in the younger generation suggests that there may be a shift towards stabilization of these tonal patterns. However, as /-h/ and /-V:T/ are each represented by only two words each and /-ʔ/ and /-VT/ by only four words each, these results may not necessarily be representative.

3.2 PCA results

3.2.1 PCA on Social Variables

Correlations between the social variables may be found in Appendix B. The percentage of variance explained by each dimension in the PCA on the social variables is displayed in Figure 11. Dimension 1, visualized in Figure 12, captures 36.7% of the variance and is primarily contributed to by most of the sociolinguistic variables except for Kuy identity, frequency of Kuy usage with
one’s own family, and time spent away. Dimension 1 appears to capture ability and frequency of Kuy usage, alongside age, which is unsurprisingly correlated with the former two factors, given the generational shift away from using Kuy. This dimension could be called Kuy Experience. However, as the hypothesis is related to usage of a tonal language (i.e. Thai and Lao in this case), this dimension will be negated and referred to as Tonal Language Experience (TLE) for ease of interpretation.

The remaining dimensions are either primarily comprised of variables that are difficult to cohesively interpret or explain too little of the variance to justify using in the analysis, so they will not be incorporated into the analysis. However, as this study is also interested in the effect of time spent living in non-Kuy speaking areas, the factor √Years Away will be employed in the analysis and will be referred to simply as Time Away.

3.2.2 PCA on acoustic correlates of voice quality

Correlations between the social variables may be found in Appendix B. The percentage of variance explained by each dimension in the PCA on the voice quality variables are displayed in Figure 13. Dimension 1 captures 41.86% of the variance and is primarily contributed to by all $H_1^*$-related values except for $H_1^* - H_2^*$ (visualized in Figure 14). This variable will be called $H_1^*(-A_n)$. The remaining dimensions will not be used as Dimension 2 is comprised mainly of CPP, Dimension 3 of $H_1^* - H_2^*$, and the rest account for little variation.
3.3 Linear Model Results

Values for the dependent variable for each register at twenty time intervals over the vowel, given values of Tonal Language Experience and Time Away 1.5 standard deviations above and below the mean, were estimated using the effects package in R (Fox & Hong, 2009). In order for readers to grasp the magnitudes of the differences, the estimated scaled values were converted back to the original unit by using raw group means and standard deviations for each gender. The back-converted values are calculated by multiplying the estimated scaled value by two times the group standard deviation and adding the group mean. Semitones are converted to Hz. These values are visualized in Figure 15. $H_{1*}$($-A_{n*}$) is not reconverted as it is a principal component comprising multiple factors but kept in “half standard deviation” units, as per the scaling method in Gelman (2008). F1 is also displayed in scaled units as the values differ by height. The linear regression table for each dependent variable may be found in Appendix C. Table 15 quantifies the mean difference between modal and breathy trajectories for each measure in Figure 15. These mean differences will be referred to as F for female participants and M for male participants.

Differences will be discussed for each measure in turn. Each figure may be thought of as representing four theoretical speakers of the stated gender who would have a TLE score 1.5 standard deviations below or above the mean and who would have spent time away equivalent to 1.5 standard deviations below or above the mean: (1) The lower left speaker, to be called the “conservative speaker”, uses Kuy, with respect to Thai/Lao, more frequently and/or proficiently than other members of the community and has spent little to no time away from home. (2) The upper left speaker has also spent little to no time away from home, but uses Kuy, with respect to Thai/Lao, less frequently and/or proficiently than other members of the community. (3) The lower right speaker uses Kuy, with respect to Thai/Lao, less frequently and/or proficiently than other members of the community, but has spent much time away from home. (4) Finally, the upper right speaker both uses Kuy, with respect to Thai/Lao, less frequently and/or proficiently than other members of the community and has also spent much time away from home.

3.3.1 f0

For both female and male speakers, less Tonal Language Experience (TLE) and less Time Away yield the smallest f0 differences between modal and breathy voice, although they are still significantly different in the expected direction (F: 3.37 Hz; M: 4.2 Hz). For female speakers, greater TLE
Figure 15: Estimated trajectories for female (left) and male (right) speakers by acoustic measure
Table 15: Mean differences between modal and breathy trajectories for each combination

<table>
<thead>
<tr>
<th>Measure</th>
<th>Vowel Height</th>
<th>Female</th>
<th>Female</th>
<th>Male</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>↑Away</td>
<td>↑Away</td>
<td>↓Away</td>
<td>↓Away</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Units</td>
<td>Scaled</td>
<td>Units</td>
<td>Scaled</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Units</td>
<td>Scaled</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Units</td>
<td>Scaled</td>
</tr>
<tr>
<td>f0</td>
<td>↑TLE</td>
<td>13.1 Hz</td>
<td>.47</td>
<td>12.55 Hz</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>3.37 Hz</td>
<td>.12</td>
<td>5.42 Hz</td>
<td>.19</td>
</tr>
<tr>
<td>CPP</td>
<td>↑TLE</td>
<td>2.40 dB</td>
<td>.33</td>
<td>1.46 dB</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>1.01 dB</td>
<td>.14</td>
<td>1.07 dB</td>
<td>.15</td>
</tr>
<tr>
<td>H₁* - H₂*</td>
<td>↑TLE</td>
<td>-1.13 dB</td>
<td>.13</td>
<td>0.43 dB</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>-1.69 dB</td>
<td>.2</td>
<td>-1.95 dB</td>
<td>.23</td>
</tr>
<tr>
<td>H₁* (- Aₙ*)</td>
<td>↑TLE</td>
<td>-.26</td>
<td>-.14</td>
<td>-.5</td>
<td>-.25</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>-.3</td>
<td>-.36</td>
<td>-.27</td>
<td>-.32</td>
</tr>
<tr>
<td>F1</td>
<td>high</td>
<td>36.68 Hz</td>
<td>.41</td>
<td>27.96 Hz</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>28.27 Hz</td>
<td>.31</td>
<td>26.56 Hz</td>
<td>.3</td>
</tr>
<tr>
<td>mid-</td>
<td>↑TLE</td>
<td>48.56 Hz</td>
<td>.31</td>
<td>35.19 Hz</td>
<td>.3</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>37.5 Hz</td>
<td>.31</td>
<td>41.55 Hz</td>
<td>.48</td>
</tr>
<tr>
<td>low</td>
<td>↑TLE</td>
<td>53.49 Hz</td>
<td>.31</td>
<td>40.81 Hz</td>
<td>.3</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>41.16 Hz</td>
<td>.31</td>
<td>47.83 Hz</td>
<td>.48</td>
</tr>
<tr>
<td>low</td>
<td>↑TLE</td>
<td>100.84 Hz</td>
<td>.31</td>
<td>76.79 Hz</td>
<td>.3</td>
</tr>
<tr>
<td></td>
<td>↓TLE</td>
<td>77.83 Hz</td>
<td>.31</td>
<td>73.06 Hz</td>
<td>.3</td>
</tr>
</tbody>
</table>

is correlated with a notable increase in f0 differences but more Time Away is only correlated with larger f0 differences in the speaker with lower TLE. Meanwhile, for male speakers, greater Time Away is correlated with a modest increase in f0 differences. The largest f0 difference between the registers is seen in the greater TLE, less Time Away female speaker at 13.1 Hz. This value is 9.73 Hz greater than the one for the conservative speaker. The largest difference for male speakers is seen in the lower TLE, greater Time Away combination at 6.06 Hz, which is 1.86 Hz greater than for the conservative speaker. A noteworthy effect is that greater TLE correlates with the f0 differences between the registers being maintained over more (and as much as all) of the vowel. More Time Away also appears to correlate with increased duration of differences, if TLE is low.

### 3.3.2 CPP

CPP shows a similar pattern to f0 in that the smallest differences are seen with less TLE and less Time Away (F: 1.01 dB; M: 0.55 dB). For female speakers, an increase of TLE correlates with greater CPP differences between the registers. For both female and male speakers, the greater TLE, less Time Away combination yields the largest differences (F: 2.4 dB, M: 2.63 dB), with the differences persisting over most or all of the whole vowel, like f0. These values are 1.39 dB and 2.08 dB greater than the values for the conservative speakers. Time Away shows mixed differences.

### 3.3.3 H₁* - H₂*

H₁* - H₂* differences are smaller for female speakers with greater TLE than those with less. The trajectory differences are virtually indiscernible for the greater TLE, greater Time Away speaker (in fact, the mean difference is positive, but only modestly so, at 0.43 dB). Male speakers, however, show the opposite pattern from female speakers, as greater TLE speakers actually have larger differences. The effect of Time Away is mixed.
3.3.4 $H_1^*(A_n)^*$

For female speakers, the largest $H_1^*(-A_n^*)$ differences are seen in the less TLE, greater Time Away combination (-0.36 scaled units) but for male speakers, they are in the greater TLE, less Time Away combination (-0.5 scaled units). An increase in TLE corresponds to a decrease of differences between the registers for female speakers, but there is no clear pattern for male speakers.

3.3.5 F1

F1 differences between the registers will be discussed in terms of scaled units, as these are the same regardless of height. The magnitude of differences for the female speaker who has spent less Time Away but has greater TLE stands out as different, being notably larger (0.41 scaled units). Notably, the increase in difference for speakers with more TLE occurs at the beginning of the vowel. Male speakers, however, show a decrease in F1 differences with both greater TLE and more Time Away. Differences between the registers are virtually indiscernible in the greater TLE, less Time Away combination (0.05 scaled units).

3.3.6 Summary of group differences

The results reveal a mixed and complicated interaction between the sociolinguistic factors of TLE and Time Away and the various correlates of the modal and breathy registers. However, a number of patterns arise.

For f0 and CPP, greater TLE shows a clear correlation with increased differences between the registers for female speakers. More Time Away is correlated with greater f0 differences for male speakers, while for female speakers, it only correlates with increased f0 differences for the low TLE speaker. Greater TLE is also correlated with the f0 difference persisting over more of the vowel for both female and male speakers. The positive correlation of TLE with f0 differences for female speakers corroborates the hypothesis, as does the positive correlation of Time Away with f0 differences for male speakers, although the magnitude is small for the latter. Greater TLE and Time Away does seem to correlate with greater persistence of f0 differences over the vowel, however. The correlation of greater TLE with larger CPP differences, at least for female speakers, however, runs contrary to the hypothesis.

Greater TLE correlates with smaller $H_1^*-H_2^*$ and $H_1^*(-A_n^*)$ differences for female speakers, while Time Away has mixed effects. For male speakers, TLE correlates with larger $H_1^*-H_2^*$ differences. The decreased $H_1^*-H_2^*$ and $H_1^*(-A_n^*)$ differences for female speakers with greater TLE align with the hypothesis, as voice quality cues are weakened. For males however, the increase of $H_1^*-H_2^*$ with greater TLE is opposite from the expectation.

Smaller F1 differences are seen with an increase in both TLE and Time Away for male speakers. For female speakers, however, there is no uniform effect, although greater TLE does seem to increase the magnitude of differences between the registers at the beginning of the vowel. The decrease for male speakers aligns with the hypothesis, while the increase of differences at the beginning of the vowel for female speakers is contrary to the hypothesis.

In sum, the hypothesis that f0 differences are increased by greater usage of Thai/Lao and more Time Away is supported by the results for male speakers, but only the effect of TLE is supported for female speakers. With respect to the effect of these factors on decreasing the weight of other voice quality measures, the results for $H_1$-related measures for female speakers and F1 for male speakers corroborate the hypothesis with respect to greater usage of Thai/Lao. Time Away also decreases F1 differences for male speakers and decreases differences in all the cues for high TLE.
female speakers. Contrary to the hypothesis, TLE is correlated with greater $H_1^*-H_2^*$ differences for male speakers and F1 differences for female speakers.

### 3.4 Individual differences

The accuracy scores of the LDA classifier for each of the five acoustic measures (f0, $H_1^*-H_2^*$, $H_1^*(-A_n^*)$, CPP, F1) and the combined VQ measure are treated as proxies for cue weights in the following analyses. A value of 0.5 for a given cue would mean that the LDA performed at chance (50%) in classifying register using only information from that cue, suggesting that the cue is completely uninformative in the contrast, while a value of 1 would mean that the LDA performed perfectly, meaning that the cue is maximally informative. Table 16 enumerates the number of speakers for which each cue is strongest, split by TLE (less TLE: < 0, more TLE: ≥ 0). It is clear from the right table that all the VQ cues combined are more informative than f0 for most speakers (59 out of 66). However, five higher TLE speakers have f0 as a more informative cue and one lower TLE male speaker shows greater informativity for F1 than VQ. With the VQ cues split up, we can see that the most important cue for lower TLE speakers in general is F1, but is f0 for female speakers and $H_1^*-H_2^*$ for male speakers with greater TLE.

Table 16: Counts for register cue with highest weight, split by TLE

<table>
<thead>
<tr>
<th>Cue</th>
<th>Less TLE</th>
<th>More TLE</th>
<th>Less TLE</th>
<th>More TLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0</td>
<td>2</td>
<td>9</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>$H_1^<em>-H_2^</em>$</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>$H_1^<em>(-A_n^</em>)$</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>CPP</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F1</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cue</th>
<th>Less TLE</th>
<th>More TLE</th>
<th>Less TLE</th>
<th>More TLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>F1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>VQ</td>
<td>15</td>
<td>14</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Three hypotheses were tested using with the results of the LDA:

1. Higher TLE and greater Time Away are correlated with stronger f0 cue weights and weaker VQ weights
2. There is a trading relation between f0 and VQ cues
3. As TLE increases, f0 increases in reliability relative to VQ as a register cue

To test hypothesis (1), Pearson’s $r$ was calculated between each register cue weight and TLE, as well as Time Away. No correlations with Time Away were significant, so these results are not displayed at all. Correlations between each cue and TLE are displayed in Table 17. Scatterplots of the significant correlations with TLE, with regression lines and a 95% confidence interval for each gender, are displayed in Figure 16. TLE is significantly positively correlated with f0 ($r = .52, p < .001$) and CPP ($r = .4, p < .05$) accuracy, and negatively correlated with $H_1^*-H_2^*$
(r = −.35, p < .05) and \(H_1^*(-A_n^*)\) (r = −.34, p < .05) accuracy for female speakers. For male speakers, TLE is significantly positively correlated with \(H_1^*-H_2^*\) (r = .5, p < .01) and VQ (r = .35, p < .05) accuracy, but negatively correlated with F1 (r = −.38, p < .05) accuracy.

Table 17: Correlations between LDA accuracy of acoustic cue and TLE

<table>
<thead>
<tr>
<th>Measure</th>
<th>(r)</th>
<th>(p)</th>
<th>Sig.</th>
<th>(r)</th>
<th>(p)</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>f0</td>
<td>.52</td>
<td>.001</td>
<td>**</td>
<td>.04</td>
<td>.838</td>
<td></td>
</tr>
<tr>
<td>(H_1^<em>-H_2^</em>)</td>
<td>-.35</td>
<td>.041</td>
<td>*</td>
<td>.5</td>
<td>.004</td>
<td>**</td>
</tr>
<tr>
<td>(H_1^<em>(-A_n^</em>))</td>
<td>-.34</td>
<td>.048</td>
<td>*</td>
<td>.05</td>
<td>.791</td>
<td></td>
</tr>
<tr>
<td>CPP</td>
<td>.4</td>
<td>.018</td>
<td>*</td>
<td>.14</td>
<td>.445</td>
<td></td>
</tr>
<tr>
<td>F1</td>
<td>.28</td>
<td>.108</td>
<td></td>
<td>-.38</td>
<td>.032</td>
<td>*</td>
</tr>
<tr>
<td>VQ</td>
<td>.01</td>
<td>.973</td>
<td></td>
<td>.35</td>
<td>.048</td>
<td></td>
</tr>
</tbody>
</table>

Figure 16: Correlation of LDA Accuracy of each acoustic measure to TLE

To test (2), Pearson’s \(r\) was first calculated between f0 and each voice quality cue, the results of which are presented in Table 18. The only significant relationship is a negative correlation between f0 and \(H_1^*-H_2^*\) for female speakers (r = −.54, p < .01). All correlations are visualized in Figure 17, with a thick black line and surrounding grey shaded region representing the regression line and a 95% confidence interval for the one significant correlation. The thin black diagonal line (\(y = x\)) is the identity line—speakers below this line have more accurate f0 scores than the compared acoustic cue, while those above the line have more accurate scores for the compared acoustic cue than for f0. One clear pattern is that female speakers under the identity line are better represented by those with greater TLE (those who lie on the red, rather than blue, end of the spectrum) in every case. Male speakers, however, have the opposite pattern when comparing
f0 to $H_1^* - H_2^*$: most red speakers lie above the identity line, while those under the identity line are almost all blue. There is also only a single male speaker, whose TLE level is average, below the identity line in the f0 to VQ comparison.

Table 18: Correlations between f0 and each voice quality cue

<table>
<thead>
<tr>
<th>Comparison</th>
<th>F</th>
<th></th>
<th>M</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r p</td>
<td></td>
<td>r p</td>
<td></td>
</tr>
<tr>
<td>f0 : $H_1^* - H_2^*$</td>
<td>-.54 .001</td>
<td>**</td>
<td>-.04 .816</td>
<td></td>
</tr>
<tr>
<td>f0 : $H_1^<em>(-A_n^</em>)$</td>
<td>-.21 .229</td>
<td>.24</td>
<td>.186</td>
<td></td>
</tr>
<tr>
<td>f0 : CPP</td>
<td>-.15 .41</td>
<td></td>
<td>.14 .452</td>
<td></td>
</tr>
<tr>
<td>f0 : F1</td>
<td>.02 .913</td>
<td></td>
<td>.29 .105</td>
<td></td>
</tr>
<tr>
<td>f0 : VQ</td>
<td>-.28 .11</td>
<td></td>
<td>0 .981</td>
<td></td>
</tr>
</tbody>
</table>

In order to test hypothesis (3), signed distance was calculated between each point and the identity line with the equation in (3), where $a = -1$, $b = 1$, and $c = 0$ (the coefficients defining the identity line) and $(x, y)$ define the coordinates of the point. For demonstration purposes, the dashed lines in Figure 18 visualize the distances that are calculated with this measure (simplifying from the graphs in Figure 17). The more positive the distance, the more informative the acoustic cue is over f0; the more negative, the more informative f0 is over the acoustic cue. Pearson’s $r$ was then calculated to determine the relationship between TLE and the informativeness of f0 as compared to the other register measures.

(3) Distance between point $(x, y)$ and line defined by $ax + by + c = 0$

$$\frac{ax + by + c}{\sqrt{a^2 + b^2}}$$

The correlations between the calculated distances for each measure comparison and TLE are presented in Table 19. Once again, Time Away shows no significant effects, so these values are left out. For female speakers, the distances from the identity line are significantly negatively correlated for comparisons between f0 and $H_1^* - H_2^*$ ($r = -.51, p < .01$), $H_1^*(-A_n^*)$ ($r = -.57, p < .001$), and VQ ($r = -.35, p < .05$). For male speakers, the only significant correlation is a positive one between the distance between the f0-$H_1^* - H_2^*$ identity line and TLE ($r = .39, p < .05$). Scatterplots showing the significant relationships, with regression lines and 95% confidence intervals for each gender, are presented in Figure 19. These results provide confirmation that the visual clustering of red coordinates below the identity line for female speakers in a number of the graphs in Figure 17 is indeed significant, as is the clustering of the red coordinates above the identity line for male speakers for the f0-$H_1^* - H_2^*$ comparison.

Table 19: Correlations between distance from identity line and TLE

<table>
<thead>
<tr>
<th>Comparison</th>
<th>F</th>
<th></th>
<th>M</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r p</td>
<td></td>
<td>r p</td>
<td></td>
</tr>
<tr>
<td>f0 : $H_1^* - H_2^*$</td>
<td>-.51 .002</td>
<td>**</td>
<td>.39 .029</td>
<td>*</td>
</tr>
<tr>
<td>f0 : $H_1^<em>(-A_n^</em>)$</td>
<td>-.57 &lt; .001</td>
<td>***</td>
<td>.003 .988</td>
<td></td>
</tr>
<tr>
<td>f0 : CPP</td>
<td>-.2 .255</td>
<td></td>
<td>.06 .736</td>
<td></td>
</tr>
<tr>
<td>f0 : F1</td>
<td>-.3 .08</td>
<td></td>
<td>-.33 .065</td>
<td></td>
</tr>
<tr>
<td>f0 : VQ</td>
<td>-.35 .042</td>
<td>*</td>
<td>.2 .278</td>
<td></td>
</tr>
</tbody>
</table>
Figure 17: Correlation of LDA Accuracy of each acoustic measure to TLE
The individual results provide evidence supporting all three hypotheses, but only for female speakers. Female speakers show f0 as an increasingly reliable cue over voice quality, particularly over $H_{1}^{*}-H_{2}^{*}$-related measures, with greater usage of a tonal language. $H_{1}^{*}-H_{2}^{*}$ is the only cue to show a trading relation with f0. Male speakers, on the other hand, do not behave as hypothesized. While greater TLE does weaken F1 cue weights, it also increases $H_{1}^{*}-H_{2}^{*}$ weights, and $H_{1}^{*}-H_{2}^{*}$ becomes an increasingly reliable cue over f0 with greater TLE. Neither group shows any significant relationships with Time Away on the individual level.

4 Discussion

This production study provides a detailed look into the complicated relationship between the shifting of cue weights and social factors related to language usage. The diversity of language experience in the Kuy population is mirrored in the large variation in cue usage among different speakers and female and male speakers have starkly different patterns.

4.1 Tonal language experience and the register contrast

The results for female speakers show a clear relationship between increased tonal language usage and heavier f0 and CPP cue weighting alongside weakening of $H_{1}^{*}-H_{2}^{*}$ and $H_{1}^{*}(-A_{n}^{*})$ differences from both the group results as well as in individual patterns. Furthermore, f0 also trades off with $H_{1}^{*}-H_{2}^{*}$ in individuals. Strengthening of F1 differences with increased tonal language usage is apparent only on the group level. Despite the heavier weighting of CPP in speakers with
greater tonal language experience, it is the most important cue for only one speaker. The evidence for a relationship between tonal language and f0 cue weighting for male speakers is visible on the group level, but not on the individual level. In Lau-Preechathammarach (in press), pitch halving in a number of modal voice tokens led to the unexpected finding of higher f0 in breathy voice than in modal voice in both female and male speakers with lower TLE and Time Away values. With these values corrected for, the unexpected finding of higher f0 in breathy voice in conservative speakers disappears, although these speakers still show the smallest f0 differences between the registers for both genders. Male speakers who use tonal languages more show larger f0 differences at the group level as well as longer persistence of differences over the course of the vowel, but show no clear importance for f0 at the individual level. The two relationships that do appear to hold on both the group and individual levels are the decreased weighting of F1 and unexpected increase in $H_1^*-H_2^*$ weighting with more tonal language usage.

One key point to be aware of in interpreting differences between the group and individual results is the fact that the LDA results are not informative about the duration over which acoustic differences between the registers persist, but rather about the overall magnitude of differences. The effect of both TLE and Time Away on f0 trajectories for male speakers is most visibly the persistence of differences over the duration of the vowel, rather than the magnitude. The loss of the dimension of time in the LDA analysis may thus obscure key differences, as contrasts are sensitive not only to magnitude, but also to temporal differences. A perception experiment with manipulation of both magnitude and contour of f0 differences would elucidate the extent to which speakers are sensitive to these factors in the register contrast.

While raw accuracy results provide some information, they may differ between speakers for reasons beyond the actual weight of the cue itself, such as task difficulty, for example. As such, within-speaker accuracy scores were compared to each other through distance from the identity line, a metric that attenuates the issue of interspeaker variation and also allows us to observe relative cue informativeness. Through this measure, the relationship between f0 and $H_1^*$-related cues for female speakers is particularly striking: we see that f0 becomes increasingly informative in relation to these voice quality cues as tonal language usage increases. For male speakers, however, the only relationship that emerges as significant is the increasing informativity of $H_1^*-H_2^*$ over f0 with greater tonal language experience, contrary to the hypothesis.

Particularly conspicuous in the female f0-VQ graph in Figure 17 is that even with all the voice quality cues combined, the LDA still performs less accurately than for f0 for five speakers, four of whom have more experience with tonal languages than average. Of these speakers, at the time of the study, two had recently graduated from college, one was a teacher, and one took up various odd jobs. The one speaker with less TLE than average was a farmer. Teachers and graduated students are currently or recently engrained in the Thai education system and the speaker who takes up odd jobs regularly interacts with members of nearby communities, many of whom are not Kuy, but rather Lao, Khmer, and/or Thai. This potentially provides some insight into why these speakers may have particularly high f0 cue weights in comparison to voice quality ones, but given the small numbers and the fact that one speaker is a farmer, this idea remains speculative. Only one male speaker, who is also a farmer, performs more accurately for f0 than when all voice quality cues are combined. This speaker’s f0 weight (83.39%) also happens to be the highest among all the speakers, being ~5% higher than the highest scoring female speaker and ~10% higher than the second highest scoring male speaker. However, this speaker lies just below the identity line and it seems the register contrast for this speaker is reliably cued by both f0 and other voice quality cues (overall VQ weight: 81.67%).

30
4.2 Separating age effects from tonal language experience

Given the fact that age is one of the main factors in the TLE principal component, one might wonder whether the shift towards heavier f0 cue weighting in female speakers is merely a generational change, rather than due to language usage. Unfortunately, because age is strongly tied to language usage, with the younger generation using Thai more frequently and proficiently than the older generation, these factors are difficult to tease apart. However, we can inspect variation within age groups. In Figure 20, Age is plotted against f0 accuracy for female speakers, with points shaded for TLE. Unsurprisingly, there is a strong significant negative correlation between Age and TLE ($r = -0.59, p < .001$); however, what is also noticeable is that in the 50 to 70 age group, most of the speakers above the regression line are lighter shades of blue than those below the line and there are even two who are shades of red. Meanwhile, within the 20 to 40 age group, the speaker with the lowest TLE (i.e. the one who is a shade of blue), has the second lowest f0 accuracy score (60.08%). The sample sizes are too small to claim that these patterns are necessarily meaningful, but the trend suggests that TLE may be responsible for heavier f0 cue weights separately from age. In order to tease these factors apart more, we would need to find more younger speakers with low TLE and more older speakers with high TLE.

4.3 Time away and the register contrast

Time spent away from home appears to have a small effect on male speakers at the group level, correlating with a modest increase in f0 and $H_1^*-H_2^*$ cue weights. At the individual level, however, no patterns are significant for either female or male speakers, but as mentioned before, this may be because of the lack of time course information for the cue. While TLE captures participants’ language usage at the time of the study, the Time Away variable was included in order to capture language usage over a participant’s lifetime, as Kuy usage is assumedly much lower in comparison to Thai and potentially Lao usage while away from home. One possibility for the lack of a clear pattern for the effect of Time Away is that it may be too coarse a variable to be an actual proxy for language usage over time. While it is likely that people use less Kuy while living in other places, the extent may differ by individual. For example, it may be the case that many younger people living in other places still use Kuy daily, as smartphones have vastly improved the
convenience for remaining in touch with those far away. One’s occupation while away from home also influences one’s social circles and subsequently one’s language usage. Another possibility is that Time Away has little, if any, effect on cue weighting, although this is improbable given the changes in social circles and language usage that come along with geographical relocation. Ultimately, the effect of time spent away from home must be explored in finer detail through a more careful inquiry into the exact nature of people’s language usage during the time they spend away from home.

4.4 Register cues for male speakers

The relationship between usage of tonal languages and f0 cue weights in production does not appear to bear out for male speakers. While the group results provide potential evidence aligning with the hypothesis, in that f0 differences persist longer for speakers who use tonal languages more or who have spent more time away from home, the individual results provide no evidence for such a relationship. One may wonder whether tonal language usage is coincidentally linked to f0 cue usage for female speakers. However, the results in Lau-Preechathammarach (in press) do show that male participants show a clear positive correlation between tonal language usage and f0 perception weights on the group level. In fact, f0 affects perception of the register contrast for male speakers more than for female speakers. Thus, the relationship between tonal language usage and f0 cue weighting does not seem to be coincidental. It appears that female speakers may be leading a change towards heavier f0 cue weighting in production, while male speakers may be leading it in perception.

One question that remains is why male speakers have increased $H_1^* - H_2^*$ weights with greater TLE. An important piece of this puzzle is that these shifting cue weights for the register contrast take place not in a vacuum, but alongside other changes. As mentioned in §1.2.2 and §2.2, two other changes in this speaker population are the existence of the syllabic nasal and the merger of coda /l/ and /r/. Sukkasem (2005) attributes the latter change to contact with Thai and Lao as both languages do not allow coda /r/. The loss of syllabic nasals is not described by Sukkasem, but as Thai and Lao lack syllabic nasals, it is not improbable that this change is related to contact as well. The loss of the syllabic nasal has implications for the functional load of the register contrast as it reduces sesquisyllabic words to monosyllabic ones, creating new minimal pairs. For example, if the syllabic nasal is dropped in the word /Nkɛŋ/ ‘waist’, it becomes /kɛŋ/, forming a minimal pair with /kɛŋ/ ‘side’. Sesquisyllabic has been proposed to be an intermediary step on the way to monosyllabicization (Brunelle & Pittayaporn, 2012, Pittayaporn, 2015), a process that has been cited as a catalyst for tonogenesis due to the pressure on the suprasegment to maintain contrast to compensate for the loss of segments (Matthes, 2001; Michaud, 2012; Kirby & Brunelle, 2017). While the merger of coda /l/ and /r/ does not affect syllable count, it does simplify the coda inventory, leading to fewer contrasts at the segmental level. With these changes promoting monosyllabic and fewer segmental contrasts, what results is an increased functional load on the correlates of register. If merger is to be avoided, the voice quality cue should be expected to be amplified, although there is no reason a priori to predict which acoustic cues would be enhanced. However, we might predict that heavy contact with a tonal language could tip the cues in favor of pitch.

It is clear from the results that overall, there is a redistribution of voice quality cue weights alongside greater experience with tonal languages. The simplified syllable structure and segmental inventory in many of these speakers creates pressures that may prompt greater unification in cues in the population for maximal salience. While cue weights are not transferred to f0 for male speakers, as hypothesized, we can understand the general enhancement of register cues to preserve
contrast in light of the decreasing segmental distinctions. For female speakers, the confluence of pressures on the suprasegmental feature, combined with ability in tonal Thai and/or Lao, primes stronger f0 cue weighting, but for male speakers, the pressure pushes cue enhancement in the direction of $H_1^* - H_2^*$, at least for production.

4.5 Conclusion

The key result of this study is that the usage of a tonal language has the potential to act as a catalyst for tonogenesis, contributing to the understanding of how language contact and bilingualism play a role in sound change through the sharing of cue usage. The findings are compatible with previous studies by Sukgasame (2003), Sukkasame (2004), and Abramson et al. (2004), who find shifts towards usage of f0 as a cue in the register contrast in other Ku(a)y communities. However, unlike the findings in Abramson et al. (2004), phonation cues are still robust in this community. They also mirror findings by Pratankiet (2001) and Sipipattanakun (2014) by demonstrating that usage of Kuy not only affects usage of Lao, but also vice versa. The shift in cue usage due to knowledge of another language aligns with previous literature on bilinguals’ L1 and L2 cue usage. Given the sociolinguistic entanglement of the four languages in the area, however, the different languages that Kuy individuals use cannot easily be categorized into L1, L2, etc. While Kuy was a first language for all participants in this study, it is also the case that many of them acquired other languages simultaneously from a young age, due both to the national status of Thai and the common ethnonlinguistic diversity of families and social circles. Although this study can not be cleanly classified as a study on L1 effects on L2 or vice versa, it does speak to the general diffusion of cues across languages within a bilingual speaker. With respect to the question of areal diffusion of tone, I do not take the strong view that Thai or Lao induce tone in Kuy, but rather align with Brunelle (2009) and Ratliff (2015a) in suggesting that preexisting f0 differences in the register contrast may be enhanced through the shared cue usage, thus making Kuy more “tone-prone”. This enhancement shifts the complex of register cue weights in Kuy, and while it does not necessitate the contrast to transform into a tonal one, the distancing of modal and breathy f0 distributions may cause listeners to be more likely to identify f0 as a meaningful cue, subsequently increasing the probability of tonogenesis occurring.

Given the long co-existence of Kuy with other groups and an extended period of quadrilingualism, one might wonder why such changes did not happen in the past. While I do not have a clear explanation for this, I propose that it is due to the changing nature of the quadrilingualism. Even if Kuy speakers have had knowledge of all four languages for centuries, they have not always used them all to equal extents. Given the long historical overlap of Kuy and Khmer territories, Khmer was the most influential non-Kuy language for a time, particularly during the period of the Khmer Empire (ninth to fifteenth century CE). As Khmer was not tonal, there would not have been a bias towards usage of f0 cues; rather, Khmer has been suggested as the source of initial devoicing, registrogenesis, and vowel restructuring in West Katuic languages (Diffloth 1982; Gehrmann, 2016, p. iii). As dynamics shifted, Lao became a more dominating presence as Kuy territory became part of the Lan Xang Kingdom (1353–1707). Why register did not give way to tone during that time could be due to chance, as sound change is of course not deterministic, but could also be related to how centralized the kingdom was. Næss & Jenny (2011, p. 230) point out that features of Mon spread into the speech of monolingual Burmese speakers in southern Burma, but not into that of monolingual Thai speakers in Central Thailand, despite the similar status of Mon in the two areas for at least three centuries. They attribute the lack of parallelism to the higher level of centralization and standardization in Thailand. If the Kuy area was not particularly incorporated into Lan Xang, the usage of Lao may not have been enough to precipitate change. A deeper under-
standing of the social dynamics and extent of Lao usage at the time would be key to understanding this part of the puzzle. In the present day, however, the extent of Thailand’s centralization policies are very clear, with schools playing a prominent role in the process and tipping the scales of multilingualism in favor of Thai through codification and Lao through social circles that arise in migration. While Kuy speakers have been in contact with Lao, Khmer, and Thai speakers and have been multilingual in these languages for a long time, it is the current social shifts that have pushed the tides of language usage towards greater usage of Thai in particular. The rich variation in the community offers insights into the mechanisms by which such transitioning behaviors of language usage translate into cue shifts, laying the groundwork for sound change.

Acknowledgements

I would like to extend special thanks to all the Kuy speakers I worked with in Tambon Tum, especially Thongwilai Intanai and Sidawun Chaiyapha, without whom this work would not have been possible. I am also highly indebted to my mentees: Margaret Asperheim, Bethany Auclair, Renee Cong, Nicole Kim, Jenkin Leung, Melissa Milligan, Geetanshi Sharma, Melody Tran, Stacey Vu, Crystal Wang, Ivori White, and Cynthia Zhong, for the hard work they put into annotating the production data. Thank you to Susan Lin, Andrew Garrett, Justin Davidson, Ron Sprouse, Pittayawat Pittayaporwn, Suwilai Premsrirat, Marc Brunelle, James Kirby, Ryan Gehrmann, Suja

Conflicts of interest

The author has no conflicts of interest to declare.

Notes


2 For Viet-Muong: Maspero (1912), Haudricourt (1954a), and Perlus (1998).


4 For comparative and phonetic work on register and voice quality, see Henderson (1952, p. 151); Gregerson (1976); Huffman (1976); Perlus (1976); Edmondson & Epling (2006). Note that tonal contrasts also involve a constellation of cues, including voice quality (Andruski & Ratliff, 2000; Abramson & Luangthongkum, 2009).
The number of semitones is calculated. 75 Hz is the reference Hz value (thus set to 0 semitones) from which the number of semitones is calculated. 75 Hz is the reference value for this study.

\[ \text{semitones} = 12 \times \log_2 \frac{H_z}{H_{z=0}} \]

A random intercept for Speaker was also not included, as it led to overfitting, likely due to the fact that values were already normalized by speaker.

Note, however, that there are only two short vowel pairs.
See Tạ (2021) for work on Arem (Vietic, Austroasiatic), in which only /i:/ shows differences, but it is actually breathy voice that is higher.  

While there is a slight drop when the coda is a glottal stop, this may be due to the fact that f0 tracking drops sharply when there is glottalization.  

This could suggest either that non-F1 voice quality cues may be very uninformative for this speaker to the point of giving “incorrect” information to the classifier or could be due to chance. F1 accuracy is at 69.15% while VQ accuracy is at 68.74%.

_p_ value significance: * = _p_ < .05, ** = _p_ < .01, *** = _p_ < .001

If the syllabic nasal is instead interpreted as prenasalization, then its loss would not lead to monosyllabilidad; however, it would instead reduce the onset inventory, thus still increasing functional load on voice quality.  

Gehrmann (2016) reconstructs Proto-Kuay without a register distinction, citing as evidence a Kuay variety that preserves the original voicing distinction and that has not undergone regirogenesis. It is possible that the larger variation in the most important cue in higher TLE speakers, many of whom are older, is indicative of an ongoing stabilization of the register distinction in the older generation. Notably, Keating et al. (2011) find that H1*−H2* appears to be the most important correlate of voice quality crosslinguistically, which may explain the redistribution of cues to H1*−H2* in male speakers, if this change is to be understood as the stabilization of register.

References


Esposito, C. M. & Khan, S. u. D. (2012). Contrastive breathiness across consonants and vowels: A compara-


Suraratdecha, S. (2014). Language and cultural rights in the ethnic revival movement of the Black Tai


